



## **Word Exchange at the Gates of Europe Five Millennia of Language Contact**

Hyllested, Adam

*Publication date:*  
2014

*Document version*  
Early version, also known as pre-print

*Citation for published version (APA):*  
Hyllested, A. (2014). *Word Exchange at the Gates of Europe: Five Millennia of Language Contact*. Det Humanistiske Fakultet, Københavns Universitet.  
[https://www.dropbox.com/s/5n0zqfuny4jgt80/Trykklar\\_PhD\\_Adam-TO.pdf](https://www.dropbox.com/s/5n0zqfuny4jgt80/Trykklar_PhD_Adam-TO.pdf)

## PhD thesis

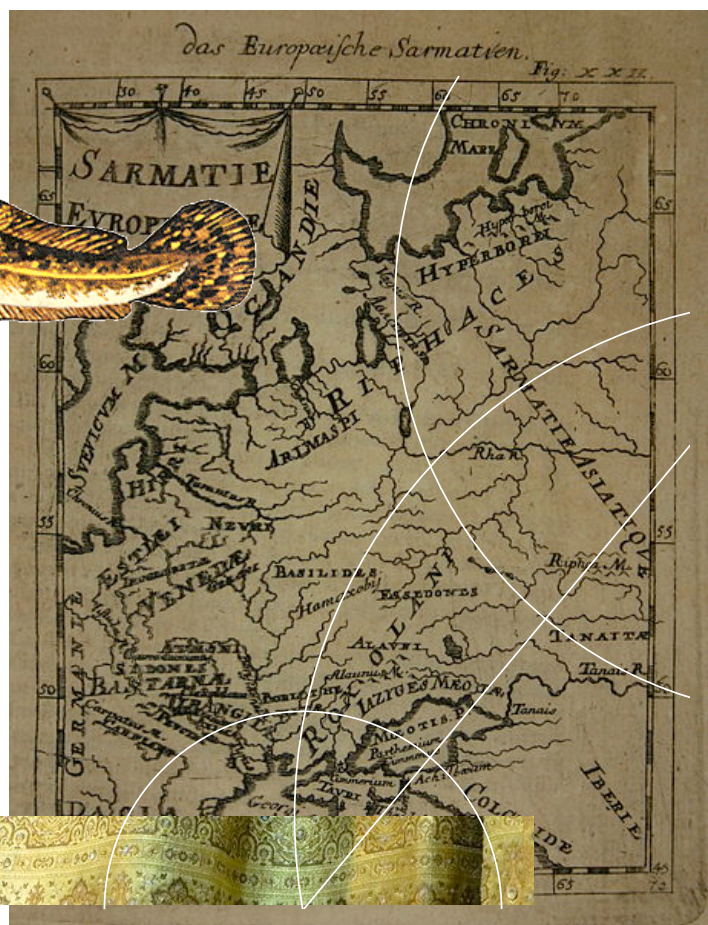
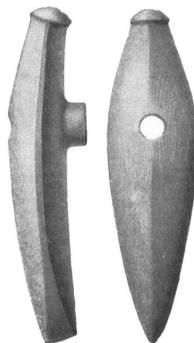
Adam Hyllested

# Word Exchange at the Gates of Europe

Five Millennia of Language Contact



S





## **Word Exchange at the Gates of Europe**

### **—Five Millennia of Language Contact**

PhD thesis  
Adam Hyllested  
University of Copenhagen, 2014

Cover illustrations:

*Sarmatie Européenne*, a map of The Grand Duchy of Lithuania by the French cartographer Alain Manesson Mallet, 1685; a Corded-Ware battle axe and hammer from Närke, Sweden; a burbot (*Lota lota*); European black elderberry (*Sambucus nigra*); fragment of a Varanasi silk sari with gold brocade; and s with *háček*, a grapheme representing an unvoiced postalveolar sibilant.



## Table of contents

Preface.....	5
Introduction.....	9
Fenno-Ugric *š as Laryngeal Substitution in Words of Indo-European Provenance .....	11
Stealing the Thunder of <i>alpaš</i> : The Fate of PIE *-bʰo- in Anatolian.....	25
The Precursors of Celtic and Germanic .....	43
Again on Pigs in Ancient Europe: The Fennic Connection .....	69
The Other Horse: Germanic Cognates of <i>caballus</i> ? .....	91
Balto-Fennic Loanwords in Proto-Germanic.....	99
Gothic <i>mammo</i> ‘meat’ in the Light of Saami Evidence .....	107
The Mysterious Elder: Common Traits in European Names for <i>Sambucus nigra</i> and <i>Viburnum opulus</i> .....	117
Balto-Fennic * <i>kakra</i> ‘oats’, the Etymology of <i>hail</i> and Another Exception to the Germanic Sound Shift .....	133
Two Issues on Indo-European Substrates in Slavic .....	153
The Story of ‘time’: The Etymology of Finnish <i>aika</i> (with an Excursus on <i>aita</i> ‘fence’ and Balto-Slavic <i>v</i> -Prothesis) .....	161
Albanian <i>hundë</i> ‘nose’, and Faroese, SWNorwegian <i>skon</i> , Finnish <i>kuono</i> ‘snout’ .....	167
<i>Estonia</i> and the <i>Aestii</i> : Baltic Etymology as a Key to Fennic Ethnonyms.....	179
Word Migration on the Silk Road: The Etymology of English <i>Silk</i> and its Congeners .....	187
English <i>mink</i> , Finnish <i>portimo</i> ‘ermine’, and Baltic Fur Trade from Antiquity to the Hanse .....	203
Latin and Slavic Loanwords in Hungarian: Exceptions to Helimski’s Vowel-harmony Adaptation Rules .....	213
English abstract .....	222
Dansk resumé .....	224

## Preface

The present dissertation is the result of my 3-year Ph.D. scholarship financed by the University of Copenhagen as a part of the five-year programme of excellence *Roots of Europe – Language, Culture, and Migrations*. It is dedicated to five people who did not live to see the final product, but whose passion for the study of language directly or indirectly continue to exercise influence on my work: First of all my enthusiastic and inspiring teacher, head of centre and initial supervisor Jens Elmegård Rasmussen whose warm and welcoming introduction to the Indo-Europeanist circles in 1994 I will never forget; my good friend, fellow etymologist and inexhaustible source of knowledge and inspiration Jan Katlev; H.E. Ambassador Paul George Jyrkänkallio, who taught me about the wonders of Hungarian grammar and passed on his personally dedicated copy of Aulis Joki's *Uralier und Indogermanen* to "min unge broder i filologi"; my grandfather, devoted Romanicist, Latinist and Germanicist, professor Povl Kristian Hyllested, whom I never got to know in person, but whose note-scribbled library I have inherited and use every day, probably the main reason I became a linguist; and my great-grandmother Ragnhild Jensen, born Høst, passionate defender and recorder of the Bornholmian dialect and penfriend in the early 20<sup>th</sup> century with U.S. ambassador to Copenhagen and Belgrade, John Dyneley Prince, who wrote, among other linguistic works, "Tatar Elements in Old Russian" (referred to in this dissertation).

I wish to thank my other teacher and final supervisor, the indefatigable Birgit Anette Rasmussen and the other colleagues at *Roots of Europe* with whom I have spent five wonderful years: Bjarne Simmelkjær Sandgaard Hansen, Anders Richardt Jørgensen, Guus Kroonen, Jenny Helena Larsson, Benedicte Nielsen Whitehead, Thomas Olander and Oliver Simkin; the tried and tested members of the *Roots of Europe* evaluation board, professors Douglas Q. Adams, Michael Janda, Joshua T. Katz, Rosemarie Lühr, Brent Vine and Andreas Willi whose advice and criti-

cal comments at their annual visits have proven invaluable; the ever-growing and talented Indo-European student community in Copenhagen, counting many dear friends; my good colleague, associate professor Tuula Eskeland who has continuously provided me with support, advice and tips on recent literature in the Uralic field; fellow Uralicist Ilda Hallas-Møller whose competent management of the library facilities has been instrumental; and a number of other people with whom I have had fruitful discussions specifically on matters treated in the final dissertation: Henning Andersen, Lars Brink, Johnny Cheung, Paul S. Cohen, Michael Fortescue, Bernd Gliwa, Berit Hildebrandt, Martin Huld, Santeri Junttila, Petri Kallio, Peter Alexander Kerkhof, Pëtr Kocharov, Agnes Korn, Kristian Kristiansen, Martin Kümmel, Ranko Matasović, Craig Melchert, Simon Mulder, Robert Orr, Kaspars Ozoliņš, Janne Saarikivi, Zsolt Simon, Merlijn de Smit, David Stifter, Patrick Stiles, Erik Thau-Knudsen, Seán Vrieland and Nicholas Zair.

This dissertation only forms part of the research I have carried out during the five project years. My employment as a PhD scholar has, in effect, been interwoven with other projects at the Roots of Europe research centre and is hard to view in isolation from them. The best known part is probably my work on the Indo-Uralic hypothesis, which was originally thought to make up the majority of my thesis. The original title of my PhD project was “The Shared Indo-European and Uralic Lexicon”, deliberately uniting the stratigraphy of loanwords on one hand and vocabulary which I believe to be inherited from a common past on the other. Scholars who investigate old loanwords tend to see them in every case of similarity, arguing vigorously against the possibility of uncovering inherited material. Adherents of the Indo-Uralic hypothesis, on the other hand, often reject rather obvious instances of borrowing, trying perhaps to maximize the amount of evidence for genetic affinity. My idea was to introduce an open mind to both approaches, including both Indo-Uralic material and older loanwords in a lexical stratigraphy. However, the Indo-Uralic part, comprising also a comparative historical phonology and derivational grammar, grew to such proportions that it will become an independent publication.

During the last half of the project I furthermore became increasingly involved in studies in Albanian language history, especially phonology, morphophonology and etymology. Finally in 2013, I devoted much time to the study of plant-names from alleged European substratum languages, a subject which ideally ought to have been included in this the-

sis but which proved to complicated for me to able to reach a satisfactory conclusion within the time-frame given.



## Word Exchange at the Gates of Europe: Five Millennia of Language Contact

### *Introduction*

The study of early lexical exchange between Indo-European and Uralic languages has a proud tradition in Denmark, not least by virtue of the the pioneering works by Vilhelm Thomsen: *Den gotiske Sprogklasses indflydelse på den finske* (1869) and *Beröringer mellem de finske og de baltiske (litavisk-lettiske) sprog: En sproghistorisk Undersøgelse* (1890). The latter still constitutes the most important reference work on contacts between Baltic and Balto-Fennic – in how many other scholarly fields today can you say that about a work written in Danish? Linguistic contacts between Uralic and Indo-European and their respective branches is still today an extensively studied and vibrant field. In some respects, though, I think that important evidence is consistently overlooked because of the power of tradition which affects not only *how* you carry out your research, but also *what* you search for, *where* you look for evidence, and *from which angle*.

Traditionally, most lexemes shared by Indo-European and Uralic language branches are viewed as having been transferred from the former to the latter. To mention the most obvious example, it is well known that both Balto-Fennic and Saami languages possess an abundance of ancient terms borrowed from (Pre-)Proto-Indo-Iranian, Proto-Baltic, Proto-Germanic, Proto-Slavic and Proto-Scandinavian. Likewise, the vocabularies of more easterly Uralic languages (Mordvin, Mari, Permian and Ob-Ugric) have been affected by intense contact with Iranian languages – apart from non-Indo-European languages such as Turkic – while Hungarian has added to its lexical stock hundreds of (Medieval) Latin, Pannonian Slavic and Alanic (Iassic) as well as Old and Middle High German loans. Samoyedic languages are even supposed to have loanwords from Proto-Tocharian. Everywhere Indo-

European is automatically assumed to be the provider and Uralic the target language while the assumed share of Uralic loanwords in older Indo-European languages is close to absent. Such an asymmetry is commonly supposed among linguists to be typical for a relationship between two peoples where one had the upper hand, technically and politically, at the time of borrowing. While the amount of borrowings rarely numbers the same on both sides, I find it unlikely that there are any cases where extensive lexical transmission in one direction leaves zero traces in the opposite direction. True, there are famous examples of extreme asymmetry such as Old Germanic languages versus Old Slavic or French versus Breton. However, even in these cases at least a small number of loans in the atypical directions are identifiable. Most such borrowings have a limited semantic and geographical distribution; they typically refer to trade objects, important plants and animals, religion, or other concepts specifically linked to the kind of contact in question.

This thesis aspires to convey to the field of Indo-European and Uralic linguistics a new methodology, where Uralic and Indo-European data are viewed as equally potential sources for loanwords. Much weight is put on our ability to reconstruct shared semantics, not least semantic anomalies, even in cases where the actual lexeme has been replaced. I also seek to underline the importance of using Uralic material as a key to unsolved issues in Indo-European. I stress especially that ignoring variation in the Uralic material (such as dialectal forms, semantic scope, irregular vocalism, forms from less well-known languages, and older attestations) can be detrimental or even fatal, leading the etymologist totally off the scent. Citing one Standard Finnish form is not enough, for example, when Balto-Fennic languages exhibit a multitude of irregular forms and deviant meanings. Finally, but equally importantly, I endeavor to establish a number of new subfields within the field of IE-Uralic contact linguistics by showing that hitherto unheeded lexical exchange took place from Proto-Mari to Proto-Baltic, from Proto-Balto-Slavic to Proto-Fenno-Permian, and from Proto-Balto-Fennic to Proto-Celtic and Proto-Germanic.

The dissertation contains 16 articles which are intended to appear in chronological order, starting with the earliest contacts.

## Fenno-Ugric \*š- as Laryngeal Substitution in Words of Indo-European Provenance<sup>1</sup>

### Abstract

Jorma Koivulehto's claim that PIE laryngeals in word-internal position are substituted with \*-š- in Fenno-Ugric must be reformulated. The development is hard to account for phonetically, and a closer look at the material reveals that there is no unambiguous evidence from outside Balto-Fennic, a subgroup of Fenno-Ugric where \*-š- regularly yields *-h-*. This means that, while word-internal laryngeals may very well be preserved in some loanwords (just like initial laryngeals are clearly preserved as *\*k-*, as shown by Koivulehto himself), they can simply have had the manifestation *\*-h-* from the beginning. If this was before the emergence of *\*h* as a phoneme, it could still have occurred as a loan phoneme. This loan phoneme would be one of the factors triggering the development of *\*-š- > -h-* in general and thereby the introduction of a new phoneme proper. One further observation is that only laryngeals before a dental stop (> BF *\*t*) are represented among the certain examples.

### 1 Frozen PIE laryngeals in loanwords in FU

In a number of ground-breaking investigations, Jorma Koivulehto (1988 [1999], 1991, 2001, 2003) has shown that in early Indo-European loanwords in Fenno-Ugric languages laryngeals are still visible, having been substituted by a variety of sounds depending on their position in the word and the time of borrowing. Most convincing is his manifestation *\*k-* for PIE *\*H-* in an older layer of loanwords, e.g.

---

<sup>1</sup> This paper was presented at the 20<sup>th</sup> Annual UCLA Indo-European Conference on 1 November, 2008.



- Fi. *kesä* ‘summer’ ← PIE *\*h<sub>1</sub>es-en-* ‘harvest’ (Koivulehto 1991: 36-40)  
 Fi. *kaski* ‘burn-beaten land’ ← PIE *\*h<sub>2</sub>ez-g<sup>(h)</sup>-* ‘ashes’ (Koivulehto 2001: 241)  
 Fi. *kalvas*, *kalpea* ‘pale’, N Saami *guolbben* ‘sandy plain; chalky layer underneath the top soil’ ← PIE *\*h<sub>2</sub>él-b<sup>h</sup>o-s* ‘white’, NHG dial. *Alven* ‘chalky layer underneath the top soil’ (Koivulehto 2003: 28)  
 N Saami *guovssu* < PSaami *\*kawsoj-* ‘dawn’ ← PIE *\*h<sub>2</sub>áus-o-* (Koivulehto 2003: 29)  
 Fenno-Volgaic *\*kuda-* ‘to weave’ ← PIE *\*h<sub>2</sub>eu-d<sup>h</sup>-* (Lith. *áudžiu* ‘id.’) (Koivulehto 1991: 50)

Koivulehto argues for a conditioning by which the velar stop *k-* occurs only in initial position, while any word-internal laryngeal (*-h<sub>1</sub>-*, *-h<sub>2</sub>-*, *-h<sub>3</sub>-*) is substituted with the postalveolar fricative *-š-* at a certain stage (but not in the oldest layer of loanwords). Fenno-Ugric *\*-š-* remains a postalveolar sibilant all the way through the Fenno-Permic, Fenno-Volgaic and Fenno-Saami (Early Proto-Fennic) stages, but develops regularly into *\*-h-* in Middle and Late Proto-Fennic, merging with the results of *\*kš*, *\*č*, *\*ž* and *\*z* (and even the original *\*-s-* of some Baltic and Germanic loanwords which may have been pronounced with a more retracted sibilant than the Fennic dental-alveolar representative). The consequence of this hypothesis is that any Finnish (or Estonian, Veps etc. – any Balto-Fennic) *-h-* may be a relic of an Indo-European laryngeal.

However, the three PIE laryngeals, although sharing notation, the term “laryngeal”, and certain behaviors such as lengthening, vowel coloring and a tendency to disappear, were clearly distinct sounds, and there is no particular reason to believe that they would all behave the same way in every position. Besides, it makes one a little bit suspicious that these postalveolar fricatives materialized as “laryngeals” (more specifically glottals) again when reaching the Balto-Fennic stage. Koivulehto in his main work on the subject (1991) gives 17 examples:

PFU or PFP form	Example	alleged NW PIE source
*(j)eškV ‘possibility’	Fi. <i>ehkä, ehki</i> ‘possibly’	* <i>ieh<sub>2</sub>-g<sup>w</sup>eh<sub>2</sub></i> (Latv. <i>jēgā</i> ‘ability’)
*(j)ešte- ‘to have enough power’	Fi. dial. <i>ehitiä</i> ‘be able to’	* <i>ieh<sub>2</sub>-g<sup>w</sup>-i%</i> - ‘be able to’
*inše ‘human being’	OFi. <i>inhe-minen</i> Fi. <i>ihminen</i>	* <i>ġenh<sub>1</sub></i> - ‘beget a child; be born’
*kešta-ta ‘to dare’	Fi. <i>kehtaa</i>	* <i>g<sup>w</sup>eh<sub>1</sub>-d<sup>h</sup></i> - ‘be ashamed’
*koneš ‘tool’ < ‘magic remedy’	Fi. <i>kone</i> ‘machine’	* <i>ġnh<sub>3</sub>-io-</i> ‘wonder’ (ON <i>kyn</i> )
*lešte ‘leaf’	Fi. <i>lehti</i> N Saami <i>lasta</i>	* <i>b<sup>h</sup>leh<sub>1</sub>-t-</i> ‘leaf; sprout’
*pewšenV ‘sieve’	Votyak <i>puž-n-</i>	* <i>peuH-eno-</i> ‘winnow- ing’
*pošta- ‘to winnow’	Fi. <i>pohta-</i>	* <i>pouH-éje-</i> ‘sift (grain)’
*püşä- ‘to fry’	Lule Saami <i>passē-</i>	* <i>b<sup>h</sup>eh<sub>1</sub></i> - ‘bake’ (OHG <i>bā(j)en</i> )
*puštas ‘clean’	Fi. <i>puhdas</i>	* <i>puH-to-s</i> ‘clean’ (Sköld 1960)
*punšV- ‘to winnow’	Mordvin <i>ponžavtoms</i>	* <i>pu-ne-H-</i> ‘cleans, win- nows’
*rešto ‘line, order’	Fi. <i>rehto</i>	* <i>(h<sub>2</sub>)ra<sub>1</sub>-to-</i> ‘line’ (Da., Sw. <i>rad</i> )
*rošto ‘grass, plant’	Fi. <i>rohto</i>	* <i>ġ<sup>h</sup>roh<sub>1</sub>-to-</i> ‘plant; growth’
*taštas ‘dough’	Fi. <i>tahdas</i>	* <i>téh<sub>2</sub>i-s-to-s</i> ‘dough’
*tešte ‘deed’	Est. <i>teht</i>	* <i>d<sup>h</sup>eh<sub>1</sub>-ti-</i> ‘deed’
*wišta ‘once; at last’	Fi. <i>vihdoin</i> ‘at last’	* <i>uiH-to-</i> ‘course, se- quence’
*wo(j)ša ‘ramification’	E Mari (Cheremis) <i>βož</i> ‘branch’	* <i>uoiH-éh<sub>2</sub></i> ‘branch’ (OI <i>vayā</i> )

Katz (2003) implicitly supported Koivulehto’s thesis regarding the reflex \*-š- although he sees them in Indo-Iranian loanwords rather than in NW Indo-European ones. We may add Mari (Cheremis) *šōžər* ‘milk’ ← PIE \**ksih<sub>2</sub>-ró-m*, which is in fact the only additional example in his book for which \*-š- explicitly substitutes *H-*.

## 2 Analyses of the individual proposals

In the following, I will comment on each of Koivulehto's relevant etymologies.

- 2.1 PFP  $*(j)E\acute{S}TE$  'TO HAVE ENOUGH POWER' ?  
 PBF  $*(j)E\acute{S}KV$  'POSSIBILITY' ← PIE  $*i\acute{E}H_2-G^wEH_2$  'POWER',  
 $*i\acute{E}H_2-G^w-i\%$  'BE ABLE TO' ?

PFP  $*(j)e\acute{s}te$  'to have enough power' (Koivulehto 1991: 77-79; > Fi. *ehtii* 'have enough time for a given purpose', dial. *ehtii olla* 'can be', N Saami *asta* 'have enough time', Mari *aste* 'do', Komi *ještj* 'be ready; be able to; be in time for; mature') is traced back to a PIE denominal verb  $*i\acute{E}H_2-g^w-i\%$  'be able to' (> Lith. *jėgti*, *jėgiù* 'be able to', Latv. *jēgt*, *jēdzu* 'id.') but this verb is actually only attested in Baltic. Because of its presence in the Permian languages it cannot be a Baltic loanword proper, at least not everywhere, and it may not even be a Balto-Slavic loanword although  $*-š-$  could reflect  $*-ž-$  as a rendering of the palatalized  $*-g'$ ; instead it could very well instead be a satem reflex of PIE  $*Hei\acute{k}$  'to have in one's power',  $*Hi\acute{k}ti$  ~  $*Ho\acute{k}ti$  (> e.g. Av. *išti* 'possession' ~ PGmc. *\*aihti* 'possession, belongings, property'), or more specifically a borrowing from Proto-Indo-Iranian into Proto-Fenno-Permic<sup>2</sup>. Phonologically, this would make the process more straightforward and account for the missing  $*j$ - in Balto-Fennic.

Koivulehto (1991: 72-74) further derives a Fenno-Saami (Early Proto-Fennic)  $*(j)e\acute{s}k\grave{a}$  'possibility' (e.g. > Fi. *ehkä* 'perhaps') from the same underlying derivative  $*i\acute{E}H_2-g^wEH_2$  'power' with reference to typological parallels like NHG *möglich* 'possible' ~ *vermögen* 'be able to'. The relevant Indo-European reflections are Lith. *pa-jėgà* 'ability', Latv. *jēga* 'id.; sensibility', Gk. *ἥβη* 'vigor; manliness; young age'. Since this word is only attested as a loanword via Proto-Balto-Fennic (Late Proto-Fennic)  $*ehk\grave{a}$ - (> e.g. Fi. *ehkä* 'perhaps'), one can further object that there is no direct evidence for a step  $-š-$  even if the Indo-European etymology should be correct; it is only reconstructed because BF  $*-h-$  is supposed to have derived from this source.

<sup>2</sup> There are related forms in Ugric (Khanty / Ostyak), but they are considered borrowings from Permian languages (SSA 100).

2.2 PFP \*KONEŠ ‘TOOL’ ← PIE \*ǵNH<sub>3</sub>-iO- ‘WONDER’?

Fi. *kone* ‘machine’ goes back to BF \**koneh* ‘tool’ which must mechanically be transposed to Early Proto-Fennic \**koneš*. Koivulehto conjectures a PIE ǵnh<sub>3</sub>-iO- ‘wonder’ as the source. Personally, I find this one of the more speculative etymologies – one might say wonderful in the sense ‘full of wonders’. More important than my own subjective impression is that -š in \**koneš* is a well-known derivational suffix in Fenno-Ugric, which may by the way sometimes represent PIE \*-s, cf. \*šommeš ‘fungus’ (PGmc. \**swambaz*), \**veneš* ‘boat’ (Skt. *vāna*- ‘timber’); so even if this is an Indo-European loan, there is no evidence as such that the laryngeals were not simply lost.

2.3 PFP \*INŠE ‘HUMAN BEING’ ← PIE \*ǵENH<sub>1</sub>- ‘BEGET A CHILD; BE BORN’?

Proto-Fenno-Volgaic \**inše* ‘human being’ is reconstructed on the basis of Mordvin (Erzya) *inže*, (Moksha) *indži* ‘guest’ and Balto-Fennic forms meaning ‘man, human being’, not least OFi. *inheminen*. However, most Finnish dialects and some of the older attestations point rather to a form \**ineh-minen*, as do all of the remaining Balto-Fennic languages, e.g. Veps *inehmoi* ‘bitch (pejorative of a woman)’, Lude *inahmoi*, Votic *inehmīn* ‘man, human being’. Although it is of course possible that a lone Old Finnish attestation could be the most archaic one, I find it more plausible that the majority of languages, dialects and old forms in this case reflect the original form. Besides, the Mordvin meaning ‘guest’ is not immediately compelling. A more probable source would then be the PGmc. antecedent of Goth. *inahs* ‘wise’ (etymologically obscure, but synchronically analyzable as *in-* + *ah-* in *aha* ‘mind’ and *ahma* ‘spirit’, an n-stem *ahman-*), either a) directly combined with the productive Balto-Fennic derivational suffix *-minen* or b) reshaped from a hypothetical \**inahman-*, consisting of the elements in both *inahs* and *ahman-*. In this case, \**inehminen* would originally have meant ‘soul’. The original -*a-* may then have been retained in Lude *inahmoi*, or the -*a-* in Goth. *inahs* may reflect PGmc. \*-*e-* (as in, e.g. *fadar* < \**faðer-*, *liuhap* < \**leuhepa-*). Another possibility is an origin in Baltic, cf. Lith. *žmonės*

‘people’, *žmonė* ‘woman’ and *i-žymi* ‘famous’ (Liukkonen 1999)<sup>3</sup>; Baltic *\*ž* is regularly substituted with Balto-Fennic *\*h*.<sup>4</sup>

#### 2.4 PFS *\*LEŠTE* ‘LEAF’ ← PIE *\*\*B<sup>h</sup>LEH<sub>1</sub>-T-* ‘LEAF; SPROUT’?

Proto-Fenno-Saami *\*lešte* can be reconstructed on the basis of PBF *\*lehte* (Fi. *lehti*, Est. *leht*) and PSaami *\*laštVš* (> N Saami *luoštaš*). Koivulehto (2003: 25), whose own surname contains a derivative of this word, *lehto* ‘grove’, sees a loan from PIE *\*b<sup>h</sup>leh<sub>1</sub>-to-(s)*, a formation reflected in PGmc. *\*blēða-* ‘blossom, sprout, fruit’ and *\*blaða-* ‘leaf’. A dedicated and distinguished Germanicist, there is no doubt that Koivulehto is happy to be able to trace a part of his surname (*koivu* means ‘birch’) back to Proto-Indo-European times, using a Germanic formation as a model for the protoform. However, *\*lehte* it is more likely to have been borrowed from some Balto-Slavic form related to Lith. *laiškas* ‘leaf’ or *lākštas* ‘sheet’, OCS *listъ* id. These forms are not completely mutually compatible except semantically, but they seem to be linked neatly by the PFS term. N Saami *lasta* cannot be from Baltic *\*lapsta* ‘leaf’ (pace Liukkonen 1999: 83–84), because *\*-a-* is substituted with N Saami *-uo-* in Baltic loans (e.g. N Saami *šuoldni* ‘dew’ ← Baltic *\*šalnā*, cf. Lith. *šalnà*).

#### 2.5 PBF *\*TAŠTAS* ‘DOUGH’ ← PIE *\*TÉH<sub>2</sub>ǵ-S-TO-S* ‘DOUGH’?

A good match both formally and semantically is PBF *\*taštas* (> Fi. *tahdas*) ‘dough’ ~ *\*téh<sub>2</sub>ǵ-s-to-s* ‘id.’ (> OCS *těsto*, OIr. *tóis*, *táis* and PGmc. *\*pāis-* in OHG *deismo* ‘sour dough’; Koivulehto 2003: 27). Crucially, however, PIE *-š-* reflects the PIE sibilant *\*-s-* rather than the laryngeal. I think this is a quite justified objection, especially since Koivulehto mysteriously does not account for the loss of the PIE *\*-s-* that is implied. Maybe he thinks that it was assimilated into the sibilant, but

<sup>3</sup> I generally do not favor most of Liukkonen’s etymologies, but this one I would definitely count as a possibility.

<sup>4</sup> Yet another possibility, if *\*inehminen* is the correct reconstruction, is to hypothesize an ancient borrowing from some IE form of *\*anə<sub>2</sub>-mo-* ‘spirit, soul’; the forms certainly match semantically and share the same skeleton of consonants, but since it is not clear precisely what language we would be talking about (we have no other evidence for a correspondence PIE *\*a-* ~ BF *\*i-*, for example), and since suffixed forms are only found in Balto-Fennic, it does not seem easy to come up with credible evidence for such a solution.

surely Ockham's razor dictates that the sibilant we encounter reflects the sibilant that was there already, while the laryngeal was lost (by assimilation or not) as laryngeals go. I have no other explanation for Koivulehto's hypothesis other than associative influence from his own native language – in Finnish the word acquires *-h-*, so when Koivulehto first saw the PIE reconstruction he may have been reminded of (and later blinded by) his own word for it. No doubt that the etymon is the correct one, though.

- 2.6 PFP \*PEWŠENV 'SIEVE' ← PIE \*PEŪH-ENO- 'WINNOWING'?,  
 PFP \*POŠTA- 'TO WINNOW' ← PIE \*POŪH-ĒġE- 'TO SIFT  
 (GRAIN)'?  
 PFB \*PUŠTAS 'CLEAN' ← PIE \*PUH-TO-S 'CLEAN'?  
 PFV \*PUNŠV- 'TO WINNOW' ← \*PU-NE-H- 'CLEANS, WIN-  
 NOWS'?

Sköld (1960: 37) was the first to point to an Indo-Iranian protoform of Ved. *pūtá-* as the source of BF \*puhta- 'clean' (> Fi. *puhdas*) and \*pohta- 'to winnow' (> Fi. *poht-i-a*). He did not visualize an intermediate step \*-š-, and it is implicit that the laryngeal was borrowed directly as \*-h-, which is also the point of this article; however, he did not give additional evidence to back this claim up. Koivulehto (1988 [1999: 301-302], 2001: 246, 2003: 26) adds PFV \*punšV- 'to winnow' (> Mordvin *ponžavtoms*) and \*pewšenV (UEW 738). These forms can hardly be separated from each other, nor from the forms mentioned by Sköld. While we are definitely dealing with a laryngeal in the root, this becomes less relevant if there are extensions with a sibilant since a sibilant in the target language is more likely to reflect that sibilant than a neighboring laryngeal. I would asserting Baltic \*pōštas 'clean' (> Lith. *puostas*) as the common source, following Liukkonen (1999: 107-108).

- 2.7 PFU \*PEŠĀ- OR \*PŪŠĀ- 'TO FRY' ← PIE \*B<sup>h</sup>EH<sub>1</sub>- 'BAKE'?

This root (Koivulehto 1981: 355-356, Koivulehto 1988 [1999: 301], Koivulehto 1991: 85, fn. 11.6) is not represented in Balto-Fennic, but is reflected in Saami (N Saami *bassi-* Lule Saami *passē-*), Permian (Udmurt/Votyak *piž-*, Komi/Zyryan *pež-*, also 'be ready/done (of food)') and Ugrian (Mansi/Vogul *pīt-*, Khanty/Ostyak *pāl-*, *pat-* 'to fry in fat').

Koivulehto asserts as the source PIE *\*b<sup>h</sup>eh<sub>1</sub>-* ‘to bake’ (φῶγω ‘to fry, to bake’, PGmc. *baka-* ‘to bake’, perhaps including the present-tense suffix *\*-i%-* (PGmc. *\*bējan-* > OHG *bā(j)en* ‘bake’)<sup>5</sup>.

Again his analysis does not address an inherent and obvious possibility that constitutes a serious weakness for the analysis of *-h<sub>1</sub>-* as the reflex of *\*-š-*: If the present suffix is potentially contained in the substituted form, where did it go? Why not just say that *\*-š-* is the rendering of both *-h<sub>1</sub>-* and the *-i-* together, probably realized as a palatal affricate? Of course the laryngeal then plays a role in the emergence of *\*-š-* in this case, but how can it then be used to explain other cases of *\*-š-* where there were no present suffixes or other *-i-*-sounds to be articulated with it. At least the loss of the suffix needs to be addressed if one (like Koivulehto) counts as a possibility that it was contained in the form that was borrowed. Note how similar this problem is to the lack of explanation of the lost *-s-* in *\*teh<sub>2</sub>istos* > *\*taštas* above – if the PIE *-s-* played a role in the emergence of PFU *-š-*, how can it be used as evidence that other laryngeals without sibilant neighbors yielded *-š-* all by themselves?

We cannot rule out that Koivulehto might be on the right track by postulating an origin in PIE *\*b<sup>h</sup>eh<sub>1</sub>-*, but in that case the origin of *\*-š-* is likely to be the PIE cluster *\*-h<sub>1</sub>-i-* and not the laryngeal alone. In my opinion, however, an origin in an unidentified satem language with secondary palatalization from PIE *\*pek<sup>w</sup>-* ‘to cook’ (Satem *\*pek-* ~ *\*peč-*, cf. Skt. *pācati*) appears more probable. In that case the attestations do not even have to reflect a single synchronic transmission, but can be the result of several independent borrowings which would explain the difficulties of reconstructing the exact vocalism. If there in fact was a PFU protoform *\*püşä-* the rounding could be explained by the preceding *\*p-*; note that PFU *\*jüwa-* ‘grain’ is from PIE *\*iēuō-* with the same kind of secondary rounding affected by the following *\*-u-* (Koivulehto 1981: 355).

My alternative proposal here obviously stands on less firm ground than those for the other items, seeing as we barely have additional evidence of the relevant correspondences in material from an early Satem dialect of Indo-European. I do think, though, that the potential existence on PFU *\*pešä-* ~ *\*püşä-* is enough reason to start searching for

<sup>5</sup> According to Koivulehto (1981: 348-356), BF *\*payi-sta-* > *\*pajsta-* (Fi. *paista-*) ‘to bake’ and BF *\*pejttä* (Fi. *peittä-*) ‘to cover’ are borrowed from each of the stems attested in Germanic, the latter via a meaning ‘wrap into something hot or something that keeps the heat’; cf. for a parallel semantic development the relationship between Olonets Karelian *suoju* ‘cover’ and Est. *sooe*, gen. *sooja* ‘heat’.

such evidence, which could also provide new important information for the discussion whether Satem isoglosses overlapped with actual PIE dialect divisions.

## 2.8 PFP \*WO(J)ša ‘BRANCH; BENDING’ ← PIE \* $\text{uoi}H\text{-}\acute{e}h_2$ ‘ID.’?

A form \**wo(j)ša* ‘branch; bending’ may be reconstructed for the Proto-Fenno-Permian stage, but in that case it disappeared completely from Fenno-Saami since it is attested only in Volgaic and Permian. It is reconstructed on the basis of Mordvin (Erzya) *užo*, Moksha *užā* ‘corner’, Mari (Cheremis) W *βaž* E *βož* ‘branch; bending, point where something bends’, and Komi (Zyryan) *vož* ‘river arm; fork; branch; (in Permyak also) sprout’. The normal reconstruction is \**woša* (Itkonen 1953-1954: 165). Koivulehto (1991: 96-99) asserts PIE \* $\text{uoi}H\text{-}\acute{e}h_2$  ‘branch’ as the source (> Ved. *vayā́*, OCS *věja* ‘id.’, derivatives of \* $\text{uei}H\text{-}$  ‘to wind’) and therefore wonders whether the FU protoform had been \**wojša*, which is also a possible reconstruction, though he prefers \**woša* after all as Balto-Fennic has no inherited words of the structure \*CVihV. He shows, however, that Udmurt (Votyak) might have preserved such a structure (cf. *vajiž* next to Komi *vož* ‘pole (on a wagon)’ < PFP \**ajša*)<sup>6</sup>, thereby providing evidence that it existed, but incidentally no reflex of the IE loan in question is attested in Udmurt. Koivulehto also thinks that it belongs to a younger layer of loanwords and therefore does not constitute an exact parallel to our ‘branch’-word.

In any case, \**wo(j)ša* is phonetically closer to PIE \* $\text{uoi}H\text{-}s\text{-}o\text{-}$  ‘branch’, yielding e.g. Slavic (LCS) \**věcha*, than \* $\text{uoi}H\text{-}\acute{e}h_2$  ‘branch’ (> OI *vayā́*, OV), incidentally an extension of the same PIE root. All other things being equal, there is of course no reason why a PIE form with *-H-* should be a more obvious source than a minimal pair with \**-s-* attested in the same daughter-languages. On the contrary, Ockham’s razor speaks for the latter. Here Koivulehto makes no methodological error, however, but simply appears unaware of the alternative IE form. Needless to say, there are no semantic problems involved in either case.

<sup>6</sup> UEW (825-826) thinks that Mordvin (Erzya) *ažija* also shows the old \**-j-*, only by metathesis, but this is rejected by Katz (1983: 118) and Koivulehto (1991: 98) who believe that the Mordvin *-j-* in this word comes from palatalization of \**-š-*.



2.9 PROTO-MARI \*šÖŽƏR ‘MILK’ ← PIE (PIIR.) \*KSIH<sub>2</sub>-RÓ-  
‘MILK’?

Now let us turn to Katz’ (2003: 193-194) lone clear example with \*-š- as a reflex of an Indo-Iranian laryngeal, Proto-Mari (Proto-Cheremis) \*šÖžƏr ‘milk’ > Meadow or Central Mari (Carevokokšaisk subdial.) šüžƏr, Hill Mari (W Mari) šəžƏr, East Mari (Malmyž subdial.) šüšƏr, Uržum Mari šör ‘id.’. He asserts PFP \*šəšƏrā ← Pre-PIIr. \*kših<sub>2</sub>róm (in his own unorthodox notation; > Ved. kṣīrá- ‘milk’, Oss.Dig. æxsir ‘id.’). This does not work either since \*šÖžƏr is not a simplex; the base-word in Mari is šör (Moiso & Saarinen 2008: 715). As shown by Aikio (2014: 131ff.), -ö- in Mari almost exclusively occurs before \*-r- or between \*n and \*l, even in loanwords.

2.10 PFS \*KEŠTA- ‘TO DARE’ ← PIE \*G<sup>w</sup>EH<sub>1</sub>-D<sup>h</sup>- ‘BE ASHAMED’?  
PFS \*REŠTO ‘LINE, ORDER’ ← PIE \*(H<sub>2</sub>)RƏ<sub>1</sub>-TO- ‘LINE’?  
PFS \*ROŠTO ‘GRASS, PLANT’ ← PIE \*G<sup>h</sup>ROH<sub>1</sub>-TO- ‘PLANT;  
GROWTH’?  
PFS \*TEŠTE ‘DEED’ ← PIE \*D<sup>h</sup>EH<sub>1</sub>-TI- ‘DEED’?

I will address the remaining examples together (Koivulehto (1988 [1999: 300-301], 2003: 26). They have probably been assigned the right Indo-European etymologies although one can always discuss the exact source language and chronological stage<sup>7</sup>. But note that they are all attested in Balto-Fennic only – where \*š yields exactly \*h! The only protoforms that we can reconstruct on the basis of direct evidence is BF \*kehta- (> Fi. *kehdata*, stem *kehta*- ‘to not be ashamed’, Est. *kõhta*- ‘to be able to’; BF \*rehto (> Fi. *rehto* ‘row, line; side; various kinds of layer’), BF \*rohto (> Fi. *rohto*, *rohtu* ‘medicine, (medicinal) plant; weed; green herb; cattle feed’, Est. *roht* ‘id.’); and BF *tehte* ‘deed’ (> Est. *teht*). Meanwhile, in all these examples the laryngeal occurs in front of a dental stop, and, correspondingly, the Balto-Fennic \*-h- precedes -t-. This means that PIE \*-h- cannot provably have been borrowed into \*-š- in Fenno-Ugric languages at all; in fact, we have no certain examples of borrowings with medial laryngeal substitutions outside Balto-Fennic. What we *can* say is that we have a handful of examples showing that a PIE *aspirated* laryn-

<sup>7</sup> Note, however, that Middle Proto-Fennic \*-kt- yields Late Proto-Fennic \*-ht-. A word like Est. *teht* ‘deed’ could therefore just as well be an inner-Balto-Fennic formation from *teke*- ‘to do’.

geal – i.e. \**h*<sub>1</sub> or \**h*<sub>2</sub> were transmitted to *Balto-Fennic* as \**h*. This is quite surprising since the loans are probably older than Balto-Fennic itself – which means that there must have existed some kind of back fricative (maybe not exactly a glottal \**h*) at least as a marginal loan phoneme in a stage before Balto-Fennic. Critics would say that this is not necessary since that stage had exactly \*š which yielded BF \**h* anyway, but there is no direct evidence for this phonologically quite odd intermediate step.

### 3 Discussion and conclusions

A more minute critical analysis of Koivulehto's and Katz's entire material remains to be carried out, but on the basis of the above considerations, I vow to conclude that, while Koivulehto's analysis that PIE initial laryngeals were substituted by \**k*- in early loanwords in Fenno-Ugric (i.e. at least at the Proto-Fenno-Ugric stage and probably even later) is virtually unchallengeable and counts as a real discovery, his bid for what happened to medial laryngeals is subject to serious misinterpretations and a high degree of uncertainty. First of all, the postulation that these three different phonemes articulated in the back of the mouth were all substituted with the postalveolar sibilant in all medial positions is hard to understand on phonological grounds. It is not so much the cross-linguistic rarity that a back fricative is replaced by a palatal fricative – we would have to accept this anyway if the material showed credible consistency – but the fact that this is combined with an unconditioned regular outcome of all three phonemes in all medial positions, independent of surroundings, makes one suspicious and demands a high standard of evidence to back it up.

This leads us to the next issue: Most of Koivulehto's etymologies involving word-internal laryngeals are not as good as those with initial laryngeals. The problem is not so much the semantics, a point where Koivulehto is markedly more cautious than many of his fellow scholars (although not necessarily a desirable approach, it is at least an uncontroversial one), but rather the fact that he in several cases overlooks alternative and more obvious candidates for Indo-European source words where \*-š- can simply be the reflex of a PIE sibilant.

A handful of Koivulehto's etymologies are serious candidates, but crucially, in these cases the only evidence comes from Balto-Fennic where \*š regularly yields *h*. This leaves us with no unambiguous piece of direct evidence for \*-š- as a manifestation of an Indo-European larynge-

al since *-h-* in Balto-Fennic can simply be direct reflexes of the similar back fricatives in Indo-European – Ockham’s razor, I would argue, dictates that there is no reason to go through an undocumented stage *\*-š-* especially if this stage is hard to account for on phonological grounds.

Except there might be one reason: It is true that *\*-h-* is normally regarded to have arisen in Middle Proto-Fennic when a postalveolar sibilant, sequences of stop + such sibilant, the voiced sibilant (in loanwords only) and the postalveolar affricate all merged into *\*-h-*. My critics may then object that there was no *\*h* present with which the IE laryngeals could be substituted because it did not exist as a phoneme before the Middle-Proto-Fennic stage – and at that stage the laryngeals are normally supposed to have been long gone from Indo-European. However, we must remember that such thing exists as a *loan* phoneme, widespread in the world’s languages. It is directly observable today that most languages, when adopting loanwords, preserve some foreign sounds that do not otherwise occur in the system. One can mention as an example *-r-* in English loanwords in continental European languages like Danish or Dutch which retain their original *-r-* quality<sup>8</sup>. In Danish, voiced sibilants and affricates do not occur, not even in English loanwords, and postalveolar sibilants are commonly substituted with the more fronted, almost palatal, variants, found in Danish native words. Thus, [ʒ] in English loanwords will be rendered by Danes typically as [ɕ]. The fact that *r* keeps its English pronunciation in the target language sets it apart as a loan phoneme. An example of a loan phoneme which has fully integrated into the system is /ʒ/ in English itself, originally from French loanwords. In fact, a commonly accepted (and inherently uncontroversial) hypothesis in Uralistics is that exactly the adoption of loanwords with *\*-š-* may have triggered the emergence of that phoneme in Proto-Fenno-Ugric. Perhaps, then, the regular rendering of a laryngeal as [h] in stages older than Balto-Fennic became one of the triggering factors that ultimately made Fenno-Saami *\*š* (as well as *\*č* and *\*kš*, and in loanwords even *\*z* and *\*ž*) develop into *\*h*, creating a fully integrated phoneme.

However, I must also conclude that we do not know at this point what the substitution was other than before a dental stop, neither what it was in Fenno-Ugric languages other than Balto-Fennic, nor what the substitution of *\*-h<sub>3</sub>-* was in any of the languages. The quest continues.

<sup>8</sup> Of course it is subject to linguists’ interpretation in every individual case when these loan phonemes start counting as real phonemes.

Koivulehto's important conclusions that a) PIE  $*h_1$  was an aspirated fricative (not a glottal stop), and b) that laryngeals were still around even at the time of NW PIE-PF(P) contacts, remain untouched.

## References

- Aikio 2014 = Luobbal Sámmol Sámmol Ánte (Ante Aikio): "On the reconstruction of Proto-Mari Vocalism". – *Journal of Language Relationship* 11: 125-157.
- Itkonen, Erkki, 1953-1954: "Zur Geschichte des Vokalismus der ersten Silbe im Tscheremissischen und in den permischen Sprachen. – *Finnisch-Ugrische Forschungen* 31: 139-345.
- Katz, Hartmut, 1983: "Hethitisch *hišša*- und Zubehör". – *Orient* 52: 116-122.
- Katz, Hartmut, 2003: *Studien zu den älteren Indoiranischen Lehnwörtern in den Uralischen Sprachen*. Heidelberg: Winter.
- Koivulehto, Jorma, 1981: "Reflexe des germ. /e<sup>1</sup>/ im Finnischen und die Datierung der germanisch-finnischen Lehnbeziehungen. – *Beiträge zur Geschichte der deutschen Sprache und Literatur* 103: 167-203, 333-376.
- Koivulehto, Jorma, 1988: "Idg. Laryngale und die finnisch-ugrische Evidenz". A. Bammesberger (ed.): *Die Laryngalthorie und die Rekonstruktion des indogermanischen Laut- und Formensystems*. Heidelberg. Pp. 281-297 [Reprinted with postscript in *Verba Mutuata*, Helsinki: Suomalais-Ugrilainen Seura, 1999, pp. 295-308].
- Koivulehto, Jorma, 1991: *Uralische Evidenz für die Laryngalthorie* [= *SbÖAW* 566]. Vienn: Verlag der Österreichische Akademie der Wissenschaften.
- Koivulehto, Jorma, 2001: "The Earliest Contacts between Indo-European and Uralic Speakers in the Light of Lexical Loans". Chr. Carpelan & al.: *Early Contacts between Uralic and Indo-European*. Helsinki. Pp. 235-263.
- Koivulehto, Jorma, 2003: "Frühe Kontakten zwischen Uralisch und Indogermanisch im nordwestindogermanischen Raum". A. Bammesberger & T. Vennemann: *Languages in Prehistoric Europe*. Heidelberg. Pp. 279-317.
- Liukkonen, Kari, 1999: *Baltisches im Finnischen* [= *Mémoires de la Société Finno-Ougrienne* 235]. Helsinki: Suomalais-Ugrilainen Seura.
- Moiso, Arto & Sirkka Saarinen, 2008: *Tscheremissisches Wörterbuch* [= *Lexica Societatis Fenno-Ugricae XXXII*]. Helsinki: Suomalais-Ugrilainen Seura.
- Sköld, Tryggve, 1960: "Drei finnische Wörter und die Laryngalthorie". – *Zeitschrift für Vergleichende Sprachforschung* LXXVI, 1/2: 27-42.
- SSA = Ulla-Maija Kulonen & al. (eds.): *Suomen Sanojen Alkuperä. Etymologinen Sanakirja*. I-III. Helsinki: Suomalaisen Kirjallisuuden Seura and Kotimaisten Kielten Tutkimuskeskus. 1995.



## Stealing the Thunder of *alpaš*: The Fate of PIE *\*-b<sup>h</sup>o-* in Anatolian<sup>1</sup>

### Abstract

Finnish *kalvas* and *kalpea* ‘pale’ and N Saami *guolbben* ‘white layer underneath the top soil’ must be Indo-European loanwords as shown by Jorma Koivulehto, indicating by their initial *\*k-* that the PIE word for ‘white’, *\*h<sub>2</sub>elb<sup>h</sup>os*, had an initial laryngeal. Kortlandt’s arguments for deeming the troublesome Hitt. *alpaš* a loanword is supplemented with the surprising fact that both the PIE nominal suffix *\*-b<sup>h</sup>o-* and the verbal root extension *\*-b<sup>h</sup>-* are virtually absent in Anatolian. Even PIE roots of the structure *CVRb<sup>h</sup>-*, of which some might at least be candidates for roots containing original *\*-b<sup>h</sup>-* extensions, turn out to be restricted to one or two examples. This remarkable state of affairs strengthens the hypothesis (presented in Hyllested 2010) that nominal *\*-b<sup>h</sup>o-* and verbal *\*-b<sup>h</sup>-* are ultimately identical. It is clear from the material that the use of *\*-b<sup>h</sup>o-* was already declining in PIE, gradually becoming replaced by other suffixes such as *\*-nt-* for the present participle. Since Anatolian was the first branch to split off the IE core, it is logical if use of *\*-b<sup>h</sup>-* was weakened further in this branch, paving the way for the multifunctional Anatolian *\*-nt-* that we know so well. It is only to be expected that a few lexicalized forms with *\*-b<sup>h</sup>o-* be preserved in Anatolian as relics, but there are simply no unambiguous examples.

---

<sup>1</sup> This paper was presented at the XII. *Arbeitstagung* of the Indo-European Society in Erlangen, September 2011, and is planned to appear as an article in *Münchener Beiträge zur Sprachwissenschaft*. However, there is still time for revision and elaboration.

# 1 The distribution, function and origin of PIE *\*-b<sup>h</sup>o-*

In an earlier article (Hyllested 2010), I have contested the widely-held views that the PIE nominal suffix *\*-b<sup>h</sup>o-* was used mainly in the formation of a) animal names, b) color adjectives, and c) abstract nouns; and that it originated by thematizations of the verbal roots *\*b<sup>h</sup>eh<sub>2</sub>-* ‘shine’ and *\*b<sup>h</sup>ueh<sub>2</sub>-* ‘be ; grow’<sup>2</sup>.

I pointed out<sup>3</sup> that animal names with *\*-b<sup>h</sup>o-* hardly ever denote the same animal in two branches, and that they are virtually all secondary formations in Indic and Greek where the suffix has been added to a vocalic or nasal stem.<sup>4</sup> None of them can be safely reconstructed for PIE, with the possible lone exception of *\*h<sub>1</sub>él-ŋ-b<sup>h</sup>o-* (and this only if Gk. ἔλαφος ‘deer’ is in fact related to PGmc. *\*lambaz/\*lambiz-* ‘sheep’ and/or the Gaulish month-name *Elembiu*).

As for the color adjectives, there are no more than four safe examples, some of which even have a limited geographical distribution<sup>5</sup>. Correspondingly, at least 20 PIE colour adjectives occur without a single attested *\*-b<sup>h</sup>-* added directly to the root in any language.<sup>6</sup> Thus, *\*-b<sup>h</sup>o-*

<sup>2</sup> The identification of *\*-b<sup>h</sup>o-* with the root *\*b<sup>h</sup>eh<sub>2</sub>-* originates from Brugmann Grdr. Bammesberger suggested that *\*b<sup>h</sup>eh<sub>2</sub>-* forms the basis of *\*-b<sup>h</sup>o-* only in color adjectives, while verbal abstracts would have *\*-b<sup>h</sup>o-* from *\*b<sup>h</sup>ueh<sub>2</sub>-*. For a modernized version of this view, see Balles 2010.

<sup>3</sup> For details and examples, I refer to the original article.

<sup>4</sup> In Old Indic, *-(a)bhá-*, in most cases with *-a-* from PIE *\*-ŋ-*, became productive in the formation of animal names, e.g. *rāsa-bha-* ‘donkey’, *śara-bhá-* ‘grasshopper’. In Greek, both the conglomerates *-a-φο-* (< *\*-ŋ-b<sup>h</sup>ó-*), *-υ-φο-*, *-ι-φο-* and the diminutives *-ά-φι-ον*, *-ύ-φι-ον* became productive and were by no means restricted to animal names; *κίρ-a-φος* ‘fox’ (~ *κιρρός* ‘orangy’); *ἄσκαλ-α-φος* ‘underworld demon, an owl’ (~ *ἄσκαλος* ‘unhoed’); *κόσσ-υ-φος* ‘blackbird’ (~ SCr. *kos* ‘id.’), *θηρ-ά-φι-ον* ‘little animal’ (~ *θήρ* ‘animal, beast’), *χωρ-ά-φ-ιον* ‘little place’, *δωρ-ύ-φ-ιον* ‘little present’; *σκίραφος* ‘trickery, cheating; gambler; dice-box’; and *ἔδαφος* ‘ground’, perhaps from *\*ued-* ‘water’.

<sup>5</sup> These are *\*d<sup>h</sup>elH-b<sup>h</sup>o-* ‘yellow’ (Arm. *deṭb* ‘yellow, blond’ ~ Arm. *delin* ‘yellow, wan, pallid’, Lat. *fulvus* ‘dark yellow’, Early Mo.Du. *dēluw*, *delluw* ‘light yellow, yellowish pale, fallow, fallow’; Driessen 2005 : 58); *\*roj-b<sup>h</sup>o-* ‘striped, spotted’ (Lith. *raĩbas* ‘grey-spotted’, OPr. *roaban* ‘striped’ ~ Lith. *raĩnas* ‘grey-spotted, striped’, OE *rāha*, *rāge* ‘roe-deer’); *\*s(u)or-b<sup>h</sup>o-* ‘dark red or black’ (Lat. *sorbum* ‘serviceberry’, OIr. *sorb* ‘stain, dirt’, Lith. *serbentà* ‘red currant’ ~ Latv. *sārts* ‘red in the face’) and *\*h<sub>2</sub>el-b<sup>h</sup>o-* ‘white’ (Gk. *ἀλφός* ‘blister’, Lat. *albus* ‘white’, PGerm. *\*alba-* ‘chalky soil’ [> NHG *Albe(n)*, Dan. *alver*], Hitt. *alpaš* ‘cloud’ ~ Lith. *alvas* ‘tin’, OHG *alunt* ‘roach’, *alant* ‘elecampane’).

<sup>6</sup> Apart from 18 roots mentioned in IEW and M&A (Hyllested 2010: 210-211), this list also includes PIE *\*d<sup>h</sup>eud<sup>h</sup>-* ‘brown’ (Skt. *dúdhitá-* ‘epithet to *tāmas-*’, Gk.

cannot have been used specifically for *color* adjectives, neither in PIE<sup>7</sup>, nor in the history of the individual IE branches<sup>8</sup>. The use of \*-b<sup>h</sup>o- in preference to other adjectival suffixes was not governed by semantics; rather, morphophonotactic restrictions seem to have applied:

- a) It occurs almost exclusively with roots ending in a sonorant, and no roots ending in a stop or -s- form adjectives with \*-b<sup>h</sup>o- added directly to the root. The other adjective suffixes \*-ro- and \*-u- are, conversely, nearly always added to obstruents (the internal distribution of \*-ro- and \*-u- being dependent on the root's syllable peak; see Rasmussen 2010). Roots ending in a laryngeal can apparently take either ending.<sup>9</sup>

---

τεῦθος 'squid', PGmc. \**duðra-* 'yellow; dodder [*Cuscuta europaea*], Toch. B *tute* 'yellow'; Schindler 1967) and PIE \*(s)*leih<sub>3</sub>-* 'blue' (IEW 965, M&A 246 'plum-coloured' + Gk. λωτός 'lotus; jujube; black tree', Skt. *nīla-* 'blue; sapphire, fig'; Hyllested 2004a + Lith. *láišis* 'tufted vetch [*Vicia cracca*]', *laišys* 'dog's mercury [*Mercurialis perennis*]', Gliwa & Hyllested 2006).

<sup>7</sup> such as \**kojH-b<sup>h</sup>o-* 'swift' > Ved. *śibhām* adv. 'fast', OHG *heif-tīg* ~ *šighrá-* 'swift', OE *hīgian* 'strive for'; \**lojh<sub>1</sub>-b<sup>h</sup>o-* 'weak' > Lith. *láibas*, *liebas* 'thin, lean', OS *lēf* ~ Lith. *leilas*, *láinas* 'thin', ON *linr* 'weak, lean', Gk. λιμός 'hunger'; and and \**nój-b<sup>h</sup>o-* > OPers. *naiba-* 'good', OIr. *nóib* 'sacred' ~ Lat. *niteō* 'shine', MIr. *nía* 'hero').

<sup>8</sup> E.g. OI *sthūla-bhá-* 'thick' ~ *sthūlá-* 'id.'; Gk. στéριφος 'stiff, hard; infertile' ~ στερεός, ON *starr* 'stiff'; and Alb. *n-gjel-bë-të* 'salty' (< \**en-sal-b<sup>h</sup>o-*). In Hyllested 2010, I included the example PGmc. \**hal-ba-* 'half' (< \**kol-b<sup>h</sup>o-* ~ Lith. *šalis* 'side'). The connection with Baltic is, however, uncertain; it is not favored by Orel 154 and not even mentioned as a possibility in Kroonen forthcoming, and even if the connection is correct, there are no obvious candidates for cognates outside the Northern European branches. Nonetheless, no alternative etymologies are generally accepted. I now believe that PGmc. \**halba-* is a loan from Fennic, cf. Fennic \**halpa-* 'reduced', gen. \**halβan-*, which cannot be a loan the other way round because the Fennic word comes from \**šalV* 'cheap'. The motivation for borrowing a word for 'reduced, cheap' would exactly be natural in a trade context which the Fennic meanings point to. Later, the Indo-European term for 'half' would have been replaced by a semantically bleached and functionally strengthened version of the Fennic word.

<sup>9</sup> The sole exceptions are inner-Greek formations of which only λίσφος 'smooth, flat' (~ λισσός 'id.') seems to be of non-onomatopoeic character; ὄσφυς 'hip' probably does not reflect \*\**h<sub>2</sub>ost-b<sup>h</sup>u-* (~ \**h<sub>2</sub>ost-* 'bone'), but is rather to be segmented \**ō-sfu-* and contains the root in σφυρός 'ankle' and σφένδα 'hurry' < \**spéu-H/d-* 'move rapidly' (cf. the connection between Eng. *hip* and *hop*). Perhaps it is even a compound \**h<sub>2</sub>ost-spu-(H/d-)-s* of the same type as OI *aṣṭhivá(nt)*, Av. *ascuua-* 'shin-bone' < Ilr. *ast-(s)čiHua-* where the second mem-



- b) Formations with  $*-b^ho-$  directly added to roots beginning with a labial stop ( $*p-$ ,  $*b^h-$ ) are avoided. We see no formations such as  $\dagger b^her-b^ho-$  or  $\dagger pelh_3-b^ho-$  in the material.

When  $-b^ho-$  occurs in deverbal nouns, these are very often result nouns<sup>10</sup>, but may also have retained an action-noun character<sup>11</sup>, or reflect earlier agent nouns<sup>12</sup>. In some cases, the distinctions are not clear<sup>13</sup>. The multifunctional use of  $*-b^ho-$  constitutes an almost exact parallel to Mod. Eng.  $-ing$ ; thus, a form like  $*skerb^ho-$  corresponds to Eng. *cutting* which is not only a present participle, but also an adjective meaning ‘capable of or designed for cutting’, an action noun meaning ‘the act of cutting’ and a result noun meaning ‘a part cut off from a main body; a clipping’.

The use of  $*-b^ho-$  for the formation of verbal nouns was rapidly declining at the time of the dissolution of PIE and remained productive in the individual branches only when accompanied by other suffixal elements<sup>14</sup>. The oldest function of  $*-b^ho-$  was the formation of present participles, indifferent to voice—i.e., both active and passive present participles.

---

ber is  $*(s)kiHu_o-$  ‘shin-bone’ (Lubotsky 2002). Hitt. *wašpaš* ‘shroud’ (Kloekhorst 2008:  $*uos-b^ho-$ ) probably contains  $-p-$  and not  $*-b^h(o)-$ , see below.

<sup>10</sup> E.g. PIE  $*gol-b^ho-$  > Goth. *kalbo*, OE *cealf* ‘calf’, Gaul. *galba* ‘fat person’ ~ Lat. *glomus* ‘bunch, wad’; PIE  $*téuH-b^ho-$  ‘swelling’ > OIr. *túaimm* ‘heap’, Gk. *τύφη* ‘pillow covering, tick’, Lat. *tūber* ‘tumor’ ~ Skt. *tūla-* n. ‘(cotton)wad’, OHG *griubo* ‘crackling, tear strip’ < ‘tearing off’; cf. Eng. *a cutting*.

<sup>11</sup> E.g. Arm. *oľb* ‘lamentation’, Gk. *ὀλόφους* ‘id.’ (~ *ὀλοφύρομαι* ‘lament’ with secondary  $-o-$  from the synonymous *ὀλολύζω*; Olsen 1999 : 37)

<sup>12</sup> E.g. PGmc.  $*wambō-$  ‘rumen’ (Goth. *wamba* ‘paunch, womb’, Dan. *vom*, . *vām*, *vom* ‘rumen, paunch’ ~ Latv. *viņbas* pl. ‘vomit, spit’ ~  $*uem(h_1)-$  ‘vomit’. Compare Lat. *rūmen* ~ Skt. *romantha-* ‘chewing the cud’.

<sup>13</sup> E.g. Lith. *gárba*, *garbė* ‘(an) honour’ ~ *giriù* ‘(to) honour’, PIE  $*lóm-b^hah_2$  ‘female water spirit’ > Gr. *νύμφη* ‘nymph’ (> Lat. *lumpa* ‘id.’, *lympa* ‘clear water’), Skt. *Rámbhā* ‘name of a water nymph’ ~ Lith. *Laumė* ‘water fairy’, older Alb. *lumet* ‘the fairies’ ~ ‘to enchant, bewitch’; Hyllested 2004b) may be an agent noun or (a concretization of) an action noun.

<sup>14</sup> For example, Baltic verbal abstracts in  $*-ī-bā-$  (Lith. *darýba* ‘building’ ~ *darýti* ‘build’; Latv. *medība* ‘hunting’ ~ *medīt* ‘hunt’; *cerība* ‘hope (subst.)’ ~ *cerēt* ‘hope (vb.)’; Lith. nominal abstracts in  $-ýbė$ , often concretised/individualised (*gražybė* ‘beauty; beautiful girl’); and Slavic verbal abstracts  $*-V-ba$  and result nouns in  $*-V-bū$  (OCS *zūloba* ‘evilness’ ~ *zūlū* ‘evil’; *gostī-ba* ‘party’ ~ *gostī* ‘guest’, SCr. *stube* ‘ladder, steps’ ~ CS *stūlati* ‘spread, stretch’).

The notion that \*-*b<sup>h</sup>o-* derives from thematicization of the verbal roots \*-*b<sup>h</sup>eh<sub>2</sub>-* ‘shine’ and \*-*b<sup>h</sup>ueh<sub>2</sub>-* ‘be ; grow’ can be easily contested on a number of grounds<sup>15</sup> :

- a) The use of the suffix is much broader than initially described by Brugmann ; thus, \*-*b<sup>h</sup>eh<sub>2</sub>-* ‘shine’ makes little sense in most adjectives in question ;
- b) The physical similarity between suffix and root is limited to a single, very frequent consonant;
- c) There are real compounds like Gk. ὑπερφύης ‘enormous, marvellous’ ~ Lat. *superbus* and Ved. *ābhva-* n. ‘monster’ < \*-*ǵ-b<sup>h</sup>u<sub>o</sub>-* (Kuiper 1962, Meier-Brügger 1991) which retain -*u-*;
- d) The alleged “modifying” effect of \*-*b<sup>h</sup>o-* (‘looking like X, of X’s kind’, ‘shining X’) lies in the modifying nature of derivation itself

The recurring extension \*-*b<sup>h</sup>-* in verbal roots (\**glei-* ~ \*-*glei-b<sup>h</sup>-* ‘smear’; \*-*h<sub>2</sub>eu-* ~ *h<sub>2</sub>u-eb<sup>h</sup>-* ‘weave’; \*-*steh<sub>2</sub>-* ~ \*-*stə<sub>2</sub>-b<sup>h</sup>-* ‘stand’ etc.) is ultimately identical to nominal \*-*b<sup>h</sup>o-* and reflect either lexicalized participial stems or simply the use of present participles for the 3rd person finite. Lexicalized or parallel formations common to IE and Uralic reveal that PIE \*-*b<sup>h</sup>o-* must be ultimately related to the PU present participial ending \*-*pa*, indifferent to voice, which is also used as the marker of the (originally unmarked) 3rd person marker of the verb.<sup>16</sup> A parallel development took place in Indo-European where another participial element, \*-*t-* ~ \*-*nt-*, came to occupy that function. Formally, then, nothing distinguishes IE verbal root variants with an extra -*b<sup>h</sup>-* from the 3rd person of the Uralic verb.

<sup>15</sup> For Slavic -*ba*, this idea was first conceived by Iljinskij (1902).

<sup>16</sup> The verb ‘to cut’, PIE \*-*sker-* and PU \*-*kere-*, provides the largest number of examples. With \*-*b<sup>h</sup>o-* and \*-*pa-*, respectively, added to the naked root, the verb ‘to cut’ acquires shared specialized meanings such as ‘to be sharp’ or ‘to scratch’, and as a noun it means ‘crust’, whereas a suffix PIE \*-*i-* ~ \*-*ja-* added before it, typically occur in derivatives denoting ‘stripes’, ‘lines’ or ‘pattern’ and in verbs meaning ‘to incite’ (later → ‘to write’). Gk. σκάριφος ‘sketch, outline; stylus’, Lat. *scribō* ‘write’, Latv. *skrīpa* ‘scratched stripes’ ~ Est. *kirjav* ‘striped, spotted’; Fi. *kirjava*; ~ *kirja* ‘pattern, figure, script’ > ‘book’). When other suffixes replace \*-*pa*, unpredictable meanings are still shared by Indo-European and Uralic (see Hyllsted 2010 for details).

## 2 Alleged examples of *\*-b<sup>h</sup>o-* in Anatolian

In any case, the use of *\*-b<sup>h</sup>o-* as a suffix must be old in PIE. It is therefore highly surprising that both nominal *\*-b<sup>h</sup>o-* and the verbal extension *\*-b<sup>h</sup>-* are virtually absent in Anatolian. In six cases, *\*-b<sup>h</sup>o-* has been suggested as the source for what seems to be a derivational ending, but they can either be refuted right away or remain disputed. Let us have a look at the candidates:

### 2.1 LYCIAN *XAĤBA* ‘GRANDCHILD’

Lyc. *xahba* (suggested by Shevoroshkin 1979: 179, fn. 5), is now known to have meant ‘grandchild’ and not ‘ruler, king’; it goes back to earlier *\*ḥaswa-* which is in itself a thematicization of *\*Honsu-* > Hitt. *ḥaššu-*, HLuw. *ḥasu-* (Melchert 1994: 63, 307).

### 2.2 LYCIAN *†xñtaba* ‘RULE’

Lyc. *†xñtaba* (suggested by Shevoroshkin 1979: 178-179) is recte *xntawa-* ‘rule’, and its source is not Luw. *ḥanda(i)-* ‘determine, fix, arrange’, which rather means ‘care for’, but the stem *ḥant-* ‘front’, cf. Luwian *ḥantawat(i)-* ‘king’ (Zsolt Simon, p.c.).

### 2.3 CARIAN *-ba-* IN PLACE-NAMES

It was suggested by Neumann (1988: 187 and n. 4) that the recurrent element *-ba-* in Carian place-names derive from PIE *\*-b<sup>h</sup>o-*. Most of the examples mentioned by Neumann are, however, etymologically quite obscure, and some of them surely continue Anatolian *\*-wa-*, e.g. *kiḍb* ‘city of Kindyē’ < *\*Hinduwa-* (cf. Simon 2008). A sound law PA *\*w* > Carian *b/C\_* can be established on the basis of *kiḍb* and *ksbo-* ‘PN’ < *\*ḥaswa-* (HLuw. *asu-*, Lycian *ḫahba-* ‘grandson; grandchild’; Simon 2008: 334).

## 2.4 HITTITE ŠALPA- ‘DOG’S EXCREMENT’

Hitt. *šalpa-* ‘dog’s excrement’ does not for sure reflect *\*sāl-b<sup>h</sup>o-* ‘grey; filthy’ (> Arm. *alb* ‘dung’) ~ *\*sal-uo-*, *\*sāl-o-* ‘dirty, grey; dirt’ (Schindler 1978; see also Olsen 1999: 37); an alternative source is still PIE *\*solp-o-* derived from *\*selp-* ‘grease; greasy’ (Ĵahowkyan 1987). If indeed derived from *\*sal-*, this item would stand alone in the sense that *\*-b<sup>h</sup>o-* would form a substantive and not an adjective. In that case, it seems appropriate either to a) reconstruct an intermediate adjective *\*sāl-b<sup>h</sup>o-* ‘dirty’ which later became substantivised or b) to assert a verbal meaning of the root ‘be dirty, produce dirt’ (cf. as a parallel the double meaning of Dan. *griset* ‘dirty, filthy’ and ‘messy’) which obtained the meaning ‘dirt’ as a kind of result noun. In any case, *šalpa-* cannot count as a safe example of a derivative with PIE *\*-b<sup>h</sup>o-*.

## 2.5 HITTITE WAŠPAŠ ‘CLOTHES (OF THE DEAD) ; SHROUD’

Hitt. *wašpaš* ‘clothes (of the dead) ; shroud’ (Goetze 1969) is related to Lat. *vespillo* ‘undertaker; grave robber’ (Watkins 1969) and derived from *\*ues-p-* ‘dress’ (see also Katz 2000). Kloekhorst (2008) reconstructs *\*uos-b<sup>h</sup>o-*, morphonotactically illicit according to me; I don’t see any reason not to accept *\*-p-* in this context. Even so, Kloekhorst might be right that we are ultimately dealing with the same morpheme – because if *\*-b<sup>h</sup>o-* does not occur following *-s-* and *\*-p-o-* does, we could argue that the extension *\*-p-* in fact reflects an allomorphic variant of *-b<sup>h</sup>-*.

## 2.6 HITTITE ALPAŠ ‘CLOUD’

Hitt. *alpaš* ‘cloud’ constitutes a problem because the expected reflex of *\*h<sub>2</sub>-* does not surface (Lubotsky 1989). Kortlandt (2003: 11) argues that it is a loanword from a non-Anatolian language. He gives five reasons:

- a) it is not found in Indo-Iranian or Tocharian
- b) it has a variant *\*elb<sup>h</sup>-* in Slavic
- c) it has an alternating suffix *-it-*, *-ut-* in Slavic and the same suffix with an infixed nasal in Slavic in the word for ‘swan’
- d) it plays a role in Germanic mythology (cf. Eng. *elf*)

- e) it is frequent in European geographical names (*Alba, Albion, Elbe, the Alps*)

We might add

- f) the fact that *-pa-* < *\*-b<sup>h</sup>o-* is virtually non-existent in the older Anatolian languages
- g) The initial stop in Fi. *kalvas, kalpea* ‘pale’ and North Saami *guolbben* (regularly < *\*kalpen-*; ~ NHG dial. *Alven*, ODan. *aluær* [mod. *al*] ‘chalky sand underneath the top soil; sandy plain’) directly reflects a laryngeal in PIE *\*h<sub>2</sub>él-b<sup>h</sup>o-* (a joint etymology by Petri Kallio and Jorma Koivulehto, see e.g. Kallio 1998 and Koivulehto 2003: 289, 298);

The seemingly absent initial laryngeal is thus secured by loans in Fennic. I do not see any semantic problem in connecting *alpaš* with Lat. *albus* ‘white’ and its cognates (as a loanword from a non-Anatolian, but still IE language), pace Puhvel HED 1/2: 38 and Kloekhorst 2008: 169; it does not *always* refer to dark thunder clouds, which can easily be covered by a generic cloud term anyway. Neither does the unique Hittite meaning need worry us since the motivation for the borrowing was most likely mythological, either going via ‘vapour, spirit’ as in Gmc. *\*albi-* ‘white creature connected with the fog’ (an original dichotomy of white *ljósal-far* as opposed to the dark *dökkálfar*, cf. also NHG *Weiße Frauen*, Dutch *Witte Wieven*) or ‘upper world’ as in Celtic, cf. also Eng. *sky* ~ ON *ský* ‘cloud’.

One might visualize a connection with Hitt. *alpant-* if this is a variant of *alwanz-* ‘being bewitched, affected by sorcery’ (Kloekhorst 2008: 171). But note that the similarity with Turkic *arba* ‘hexen, bezaubern, wahrsagen’, at first glance of course superficial, constitutes a parallel to Gk. ἄλφι n. ‘barley-groats’, Alb. *elp*, *-bi* ‘barley’ vs. Turkic *arpa* ‘barley’, Mong. *arbaj* id. The narrow semantics is in both cases coupled with a correspondence between *-l-* in the Indo-European forms and *-r-* in Altaic. At the same time, a form *\*arpa* also occurs in Uralic word for ‘withcraft’. As is well known, religious and agricultural terms are both typical loanwords. While Blažek (2012) prefers an Indo-European (Iranian) origin of the Altaic words for ‘barley’, several facts do point to a borrowing in the reverse direction. First of all, the word is “unusually common in the Turkic languages” (Stachowski 12); the Turkic word is already regarded the source of the Mongolian and Tungusic forms;

within Iranian, the word is not found outside East Iranian (Stachowski); and its only cognates are found in Greek and Albanian, both Balkan languages. As Stachowski writes, most previous works have uncritically quoted previous works about the possible Iranian origin of the Turkic term. Tatarincev (2000) suggests that the word is an inner-Turkic derivative, formed by \**ar*- ‘multiply oneself, be numerous’ with a suffix denoting intensification, cf. Old Turkic *arka* ‘multitude; collection; crowd, group’, Mong. *arbin* ‘plentiful’. Martin (1987) and Omodaka (2000) have added OJap. \**apa* ‘millet’ as a plausible cognate; I do not see how Stachowski can conclude that this speaks for Tatarincev’s inner-Turkic derivation, but in any case it strengthens the hypothesis that we are dealing with an Altaic agricultural term of great age.

If the ‘barley’-word is indeed of Turkic or even Altaic origin, it seems justified to hypothesize a similar origin of Hitt. *alpant*- ~ *alwanz*-. The lambdacization in either word does not have to have happened after the borrowing since confusion between liquids is a common phenomenon already within older Altaic languages (Granberg 2008). However, there is also a possibility that \**arpa* is a Uralic word borrowed into Turkic at an early stage if *-pa* could be identified as the participial suffix.

Whatever the exact history of these two words, most signs point to an extra-Indo-European origin of both of them. Hence, *alpant*- and *alwanz*-, as well as the designations for ‘barley’, should be kept apart from *alpaš* until stronger evidence for a connection shows up.

### 3 Verbal \*-*b<sup>h</sup>*- and roots of the structure CVRb<sup>h</sup>- in Anatolian

Let us now have a look at Anatolian verbal roots of the structure *CeR*(-)*b<sup>h</sup>*-, since these are all roots that could possibly contain a verbal extension \*-*b<sup>h</sup>*- :

#### 3.1 HITTITE *KARP-IEZZI* ‘TO TAKE (AWAY), LIFT (UP), PLUCK’

Kloekhorst (1998: 453) derives Hitt. *karp-িয়েzzi* ‘to take (away), lift up, lift, pluck’ from PIE \*(s)*kerp*- (Lat. *carpō* ‘pick, pluck’ etc.), as opposed to Oettinger (1979: 345) who traces it back to PIE \**g<sup>h</sup>reb<sup>(h)</sup>**h*<sub>1</sub>- ‘to dig’. Even in the latter case, it seems to have root-final \*-*h*<sub>1</sub>- (see Olsen 1993) and thus does not count as an original example of \*-*b<sup>h</sup>*-.

3.2 HITTITE *KARAPI*, *KARE/IPANZI* ‘TO DEVOUR’

Kloekhorst (2008: 442-444) derives Hitt. *karapi*, *kare/ipanzi* ‘to devour’ from the root *\*g<sup>h</sup>erbh<sub>1</sub>-* (> Skt. *gṛbhnāti* ‘to seize’ and ON *grápa* ‘id.’) But ON *grápa* continues < PGmc. *\*grēpP-* < PIE *\*g<sup>h</sup>rēb<sup>h</sup>-n-* rather than *\*g<sup>h</sup>rēb-* (Kroonen 2012). Instead this verb should be grouped with Nw. *garva*, *garpa*, *gurpa* ‘devour, gobble, belch’, assuming the doublets reflect an ablauting iterative doublet *\*garppōḫi*, *\*garbunanḫi* ~ *\*gurppōḫi*, *\*gurbunanḫi* < *\*g(o)rb<sup>h</sup>-néh<sub>2</sub>-ti*, *\*gh(o)rb<sup>h</sup>-nh<sub>2</sub>-énti*. In either case, the aspiration of the root-final *\*-b<sup>h</sup>-* seems not to be original, but reflects PIE *\*-b<sup>(h)</sup>-H-* (see again Olsen 1993).

3.3 HITTITE *ḪARP(P)*- ‘TO CHANGE ALLEGIANCE’

Kloekhorst (1998: 311, 2008: 442-444) connects Hitt. *ḫarp(p)*- ‘to change allegiance etc.’ to PIE *\*h<sub>3</sub>erb<sup>h</sup>-* (> Gk. ὀρφανός ‘orphan’ etc.). This is the only certain example of a final *\*-b<sup>h</sup>-* in a triconsonantal root, but it still bears no signs of having resulted from extension of a shorter root.

3.4 HITTITE *ḪUPPIYA*- ‘THROWS, HURLS’

In Hyllested & Cohen (2007), our aim was to show that it is phonologically unproblematic to link Gk. ὑφαίνω ‘weave’ to Hitt. *ḫuwapp-* (allegedly ‘interlace, entangle’, Puhvel 1991), despite the lack of prothetic vowel in Greek. This is because there are no examples of initial u-diphthongs before a labial in Greek except for late inner-Gk. formations; both full-grade *\*(H)euP-* and zero-grade *\*HuP-* regularly yield Proto-Gk. *\*uP-*. Recently, however, Melchert (2007) has shown that *ḫuwapp-* rather means ‘throw, hurl’. This obviously does not contradict the Greek rule, but it does remove one important piece of positive evidence, and, more importantly, it seems to undermine the evidence for *\*h<sub>2</sub>* in ‘weave’ (cf. also van Beek 2011).

Since Neu (1998), another candidate for a cognate of *weave* etc. has been the hapax *wepuš wēpta* in the fragment KBo. 42.6, 9 (13<sup>th</sup> c. BC) whose exact interpretation is still debatable:

(8) [...] *-zi-mi-iš* ÍD-*aš ar-ru-ma-ar e-ep-t*[*a ...*]

(9) [...] *x-ni ú-e-pu-uš ú-e-ep-ta nu=mu TÚG-an=mi-i*[*t ...*]

‘[sby] took the washing of the river [or in the rivers]  
 ‘[sby] *wep*-ed *wep*-s and [...ed] my clothing for me’

Reconstructing PIE ‘weave’ as either *\*h<sub>1</sub>ueb<sup>h</sup>-* or *\*ueb<sup>h</sup>-* is problematic since it precludes a *-b<sup>h</sup>-* extension of PIE *\*h<sub>2</sub>eu-* ‘id.’ > e.g. Skt. *váyati*, Lith. *áudžiu* (< *\*h<sub>2</sub>eu-d-*), whose initial laryngeal is needed to account for the initial *\*k-* in the Fenno-Permian loan *\*kuḍa-* ‘weave’ (> Fi. *kuto-*, N Saami *godde-*, Mordv. *koda-*, Komi *kyj-*).

Furthermore, the VN *arrumar* ‘washing’ is mentioned *before* the *wep*-sequence, suggesting this does not refer to fulling. It could refer to the washing of wool fibers before the preparation for spinning, but this was carried out in *hot* water (i.e. not a river), and the process involved a lot of intermediate activities – drying, beating, cleansing, carding, grading, bleaching – not mentioned in the fragment (Breniquet 2010).

Since the concept of ‘throwing’ is central to also to ancient weaving, cf.:

- a) Eng. *warp* ~ OE *weorpan* ‘to throw’ and
- b) *to throw the shuttle*

it could be that *ḫuwapp-* belongs with *weave* after all, having preserved an original PIE meaning that was specialized in Core IE after the Anatolian split-off; Andrés-Toledo (2010), too, suggests a late semantic narrowing, but from an original meaning ‘bind, interlace’, based partly on the now rejected Hitt. meaning, and partly on Indo-Iranian which also displays the meaning ‘weave’. As Melchert notes, the Hitt. hapax *ḫūpala-* ‘fish-net’ does not have to be derived from a verb ‘weave’ because a net is something you *cast out*.

The sumerogram TÚG ‘clothes’ represents Hitt. *wašpaš* (Goetze 1969) which often specifically means ‘shroud; clothes of the dead’, hence Lat. *vespillo* ‘undertaker; grave robber’ (Watkins 1969) < *\*ues-p-* ‘dress’ (see also Katz 2000). It occurs elsewhere in Hitt. texts that a dying man himself is calling for his shroud or his washing. In the Old Hittite-Akkadian Testament (§ 3, Kbo III 64-73, Melchert 1991: 183), the dying king Hattušiliš says ‘wash me well; protect me at your bosom from the earth’ (Melchert 1986) and the Soldier’s Dirge reads

*Nešaš wašpeš Nešaš wašpeš*  
*tiya=mmu tiya*



*nu=mu annaš=maš katta arnut*  
*tiya=mmu tiya*  
*nu=mu uwaš=maš katta arnut*  
*tiya=mmu tiya*

‘shrouds of Neša, shrouds of Neša  
 wrap me, wrap  
 put me down for burial with my mother  
 wrap me, wrap  
 put me down for burial with my forefathers  
 wrap me, wrap’

It is noteworthy that the passage makes use of alliteration (involving four words):

... *ēpta* ... *wepuš wēpta* ... \**wašpan* ...

Because a *figura etymologica* of a similar shape, again used with a word for ‘toga’ in the acc., occurs in the S Picene epitaph TE 2 from Bellante:

**postin : viam : videtas : tetis : tokam : alies : esmen : vepses : vepeten**  
 ‘along the road you see / the toga (*or* covering) of Titus Allius (?) / buried (?) in this tomb (?)

This stylistic feature is of PIE age (Watkins 1995: 131-133, Fortson 2002: 73), and the SPic. *vep-* even occurs in non-etymological alliterations such as *veiat vepetí* ‘lies ... in the tomb’ in MC 1 from Loro Piceno:

**apaes : qumat : esmín : púpúnis : nír : mefiún : veiat : vepetí**  
 ‘The elder lies, the Picene chief, in the middle of the tomb’

Correspondingly, Ved. *vap-* ‘strew’ and its derivative *vapuṣ-* ‘wonder’, transposed to meanings like ‘color; covering; clothing’, forms an alliterative pair with *vas-* ‘to wear, to dress’ (Katz 2009, Jackson):

RV 3.55 14 (Heaven and Earth according to Sāyaṇa; but perhaps Dawn and her Sun-god husband, Sūrya)

*pádyā vaste pururúpā vápūnsy ūrdhvā tasthau tryáviṃ rérihānā*

*ṛtāsya sādma ví carāmi vidvān mahād devānām asuratvām ékam*

‘Unten kleidet sich die Vielfarbige in schöne Formen; sie richtet sich empor, das anderthalbjährige Rind leckend. Ich durchwandere als Wissender die Stätte der Wahrheit. Groß ist die einzige Asuramacht der Götter.’

RV 1.160, 2 (Heaven and Earth):

*uruvyācāsā mahīnī asaścātā pitā mātā ca bhuvanāni rakṣataḥ  
sudhṛṣṭame vapusyē ná ródasī pitā yāt sīm abhī rūpaír āvāsayat*

‘Breiträumig, großmächtig, nie versiegend, behüten Vater und Mutter die Geschöpfe. Die sehr kecken (?) Rodasī sind wie zwei schöne Frauen, da der Vater sie in Farben kleidete.’

MS 3.11.9

*sārasvati manasā peśalam vasu nāsatyabhyām vayati darśatām vāpuh  
peśalam, cf. RV 1.92.4 (Dawn): *pēsānsi vapate*, but no figura etymologica (Katz 2009)*

It is thus conceivable that KBo. 42.6, 9 describes a burial rite with a dying or even dead person speaking, and that both the Hittite and South Picene items represent PIE \**uēp*- ‘to adorn, to make ready by adorning’ (*pace* Meiser 48-49). Katz (2009) adds to this root Gk. ὀπνίω ‘marry’ on the basis of a new sound-law for Greek that makes \**u*- disappear in this context.

I see no reason to leave out ORu. *vapī* ‘color’, *vapīno* ‘chalk’, OPr. *woapis* ‘color’ and Latv. *vāpe* ‘glaze’ from this family; cf. the parallel in OPr. *sirmen* ‘funeral rite’ ~ *sirmes* ‘washing lye made of ashes’ ~ Lith. *širmas* ‘white; grey’ (Gliwa 2005). On the concept of color in prehistoric funeral rites in general, see Jones-Bley (2005).

The spelling with single -*p*- in *wepuś* is problematic, but it is a hapax preceding *wēpta* which may have influenced it. The plene spelling of *ḥuwapp-* is no less problematic, but at least the semantic comparison between a verb meaning ‘to throw’ and ‘to weave’ need not be.

Katz connects them with Hitt. *wappu-* ‘riverbank’ (< ‘heaped-up earth’) and Skt. *vāpra-* ‘heap of earth via the meaning ‘heaped up (finery)’. I propose an alternative: These belong with \**ōferaz* ~ \**ōferan-*

‘bank, shore’ (MHG *uover*, NHG *Ufer* etc.), Gk. ἡπειρος ‘the land as opposed to the sea’ and perhaps Lith. *upė* ‘river’. This would be the only case of expected PGmc. *\*wōf-* ~ *\*wōb-*, and it is conceivable that such a sequence with two labial fricatives and a rounded vowel in the middle would be subject to dissimilation.

#### 4 Conclusions

The lack of evidence for both nominal *\*-b<sup>h</sup>o-* and verbal extensions in *\*-b<sup>h</sup>-* in Anatolian strengthens the hypothesis that these two elements are ultimately identical. They were not derived from verbal roots in PIE; rather do they belong to a more distant past where they formed present participles indifferent to voice (like PU *\*-pa*), and, like Eng. *-ing*, it ended up synchronically as a derivational suffix for both agent nouns, abstract nouns, result nouns and adjectives. Its occurrence in animal-names is language-specific, based on substantivizations of color adjectives. PIE *\*-b<sup>h</sup>o-*, thus already declining as a participle marker, gradually became replaced by *\*-nt-* (which also has a counterpart in Uralic). Since Anatolian broke off the core first, it is logical if the tendency was weakened further (and *\*-nt-* correspondingly strengthened) in this branch. As is well-known, the use of *\*-nt-* in Anatolian goes far beyond the formation of participles. We would expect a few lexicalized forms with *\*-b<sup>h</sup>o-* to be preserved as relics, although not necessarily for us to study as attestations in the corpus. What is relevant is not whether we can eliminate the examples altogether, but that we have so few of them in any case.

#### References

- Adiego, Ignacio J., 2007: *The Carian Language*. Leiden / Boston.  
 AiGr = Jacob Wackernagel (–Albert Debrunner et al.) (1896)/1930–1957: *Altindische Grammatik*. Vols. I–III. Göttingen: Vandenhoeck & Ruprecht.  
 Andersen, Henning, 1998: “A Glimpse of the Homeland of the Slavs: Ecological and Cultural Change in Prehistory”. – A. Della Volpe (ed.) 1998, pp. 1–67.  
 Andersen, Henning, 2003: “Slavic and the Indo-European Migrations”. – H. Andersen (ed.) 2003, pp. 45–76.  
 Andersen, Henning (ed.), 2003: *Language Contacts in Prehistory* [= *Current Issues in Linguistic Theory* 239]. Amsterdam / Philadelphia: John Benjamins.  
 Andrés-Toledo, Miguel Ángel, 2010: “Some considerations about Vedic, Avestan and Indo-iranian Textile Terminology”. – Michel & Nosch pp. 431–444.

- Balles, Irene, forthcoming: "The Old Indic *cvi* construction, the Caland system, and the PIE Adjective". – J.E. Rasmussen & T. Olander (eds.): *Internal Reconstruction in Indo-European*. Copenhagen: Museum Tusculanum Press.
- Bammesberger, Alfred, 1973: *Abstraktbildungen in den baltischen Sprachen*. Göttingen: Vandenhoeck & Ruprecht.
- Bammesberger, Alfred, 1998: "Die etymologische Herleitung von lateinisch *morbus*". – *Studia Etymologica Cracoviensia (SEC)* 3.
- van Beek, Lucien, 2011: "Vowel Assimilation in Greek: The Evidence Reconsidered". – Th. Krisch & Th. Lindner: *Indogermanistik und Linguistik im Dialog*. Wiesbaden: Reichert. Pp. 49-58.
- Benveniste, Émile, 1949: "Noms d'animaux en indo-européen". – *Bulletin de la Société Linguistique de Paris* 45, pp. 74-103 (III. "Noms du "mâle"", pp. 100-103).
- Breniquet, Catherine, 2010: "Weaving in Mesopotamia during the Bronze Age". – In Michel & Nosch pp. 52-67.
- Brugmann Grdr. = Karl Brugmann, 1897-1916: *Grundriß der vergleichenden Grammatik der idg. Sprachen* I-III. Strasbourg: Karl J. Trübner.
- Campanile, Enrico, 1983: "Le Restsprachen e la ricerca indoeuropeistica". – Edoardo Vineis (ed.), *Le lingue indoeuropee di frammentaria attestazione – Die indogermanischen Restsprachen* [= *Biblioteca della Società Italiana di Glottologia* 4]. Pisa: Giardini, pp. 211-224.
- Carpelan, Christian, Asko Parpola & Petteri Koskikallio (eds.), 2001: *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations* [= *Mémoires de la Société Finno-Ougrienne* 242]. Helsinki: Suomalais-Ugrilainen Seura.
- Chantraine DELG = Pierre Chantraine, 1968-1980: *Dictionnaire étymologique de la langue grecque*. Paris: Klincksieck.
- Čop, Bojan, 1973: *Prispevek k zgodovini labialnih pripon v indoeuropskih jezikih / Beitrag zur Geschichte der Labialsuffixe in den indogermanischen Sprachen* [= *Razred za filološke in literarne vede, Classis II: philologia et litterae, dela / opera* 29]. Ljubljana: SAZU.
- Della Volpe, Angela (ed.), in cooperation with Edgar C. Polomé, 1998: *Proceedings of the Seventh Annual UCLA Indo-European Conference 1995* [= *JIES Studies and Monograph Series* 28], Washington, D.C.: Institute for the Study of Man.
- Driessen, C. Michiel, 2005: "On the etymology of Lat. *fulvus*". – Meiser & Hackstein 2005, pp. 39-64.
- Fortson, Benjamin W., IV, 2005: "Linguistic and cultural notes on Latin *Iūnius* and related topics". – Mark R.V. Southern (ed.): *Indo-European Perspectives* [= *Journal of Indo-European Studies Monograph Series* 43]. Pp. 61-77.
- Frisk GEW = Hjalmar Frisk, 1960-1972: *Griechisches etymologisches Wörterbuch*. Heidelberg: Carl Winter Universitätsverlag.
- Gliwa, Bernd, 2005: "Baltische Bestattungsrituale im Spiegel ihre Bezeichnungen: litauisch *šermenys* und *šarvójimas*, altpreußisch *sirmen*". – *Acta Linguistica Lithuanica* 53: 9-21.
- Gliwa, Bernd, 2007: "Altpreußisch *tubo* – *filcz* (E 488)". – *Res Balticae* 11.
- Gliwa, Bernd & Adam Hyllested, 2006: "Baltisches zu Idg. *\*sleǵH-* 'blau (sein)'. – *Baltistica* XLI (1), pp. 87-93.
- Goetze, Albrecht, 1955: "Hittite Dress". – *Corolla Linguistica* [Fs. Sommer], pp. 48-62.

- Granberg, Antoaneta, 2008: "The Hunno-Bulgarian language". – *The DSCA Journal* 2008.
- Hyllested, Adam, 2004a: "Greek *λωτός* 'lotus' and the Indo-European Words for 'blue'". – James Clackson & Birgit Anette Olsen: *Indo-European Word Formation*. Copenhagen: Museum Tusculanum Press, pp. 59-64.
- Hyllested, Adam, 2004b: "L'esprit des eaux: grec *νύμφη*, sanskrit *Rāmbhā*, lituanien *Laumė* et quelques autres formes semblant apparentées". – *Per Aspera ad Asteriscos* [Fs Rasmussen]. Innsbruck: IBS, pp. 219-233.
- Hyllested, Adam, 2004c: Review of Andersen (ed.) 2003. – *Acta Linguistica Hafniensia* 36: 179-184.
- Hyllested, Adam, 2009: "PIE *\*-b<sup>h</sup>-* in Verbs and Nouns: Distribution, Function, Origin". – R. Lühr & S. Ziegler (eds.): *Protolanguage and Prehistory*. Wiesbaden: Reichert. Pp. 202-214.
- Hyllested, Adam, 2010: "Internal Derivation vs. External Comparison: The Case of the Indo-Uralic Laryngeals". – Jens Elmegård Rasmussen & Thomas Olander (eds.): *Internal Reconstruction in Indo-European*. Copenhagen: Museum Tusculanum Press.
- Hyllested, Adam & Paul S. Cohen, 2007: "Monophthong for Expected *v*-diphthong in Greek". – Coulter George, Matthew McCullagh, Benedicte Nielsen, Antonia Ruppel & Olga Tribulato (eds.), *Greek and Latin from an Indo-European Perspective* [= *Proceedings of the Cambridge Philological Society Supplementary Volume* 32]. Cambridge: Cambridge Philological Society.
- IEW = Julius Pokorny, 1959: *Indogermanisches Etymologische Wörterbuch*. Tübingen / Basel: Francke.
- Iljinskij, Grigorij, 1902: "Zur slavischen Wortbildung". – *Archiv für Slavische Philologie* 24: 224-228.
- Ġahowkian, G.B., 1987: *Hayoc' lezvi patmowt' yown*. Yerevan.
- Janda, Michael, 2001: Review of Della Volpe (ed.) 1998. – *Kratylos* 46 : 202-205.
- Janhunen, Juha, 1982: "On the Structure of Proto-Uralic". – *Finnisch-Ugrische Forschungen* 44: 23-42.
- Jokl, Norbert, 1911: "Studien zur albanesischen Etymologie und Wortbildung". – *Sitzungsbericht der Wiener Akademie der Wissenschaft Philosophisch-historischer Klasse* 168, Teil 1.
- Jones-Bley, Karlene, 2005: "Red for the Dead". – D. Disterheft, M. Huld & J. Grepin: *Studies in Honor of Jaan Puhvel* I: 211-2005.
- Kallio, Petri, 1998: "*Kalja*". – *Sananjalka* 40, pp. 87-93 (abstract: "*Kalja* 'beer'", p. 94).
- Katz, Joshua T., 2000: "Evening Dress: The Metaphorical Background of Latin *uesper* and Greek *ἑσπερος*". – K. Jones-Bley, M.E. Huld & A. Della Volpe (eds.): *Proceedings of the Eleventh Annual UCLA Indo-European Conference*. Washington D.C. Pp. 69-93.
- Katz, Joshua T., 2009: "Greek *ὀπιώ* and its digamma". Ms. / paper presented at *Indo-European Matters* in Copenhagen 9 Oct, 2009.
- Kloekhorst, Alwin, 2008: *Etymological Dictionary of the Hittite Inherited Lexicon*. Leiden: Brill.
- Koivulehto, Jorma, 2001: "The Earliest Contacts between Indo-European and Uralic Speakers in the Light of Lexical Loans". – Carpelan, Parpola & Koskikallio (eds.): pp. 235-263.

- Koivulehto, Jorma, 2003: "Frühe Kontakte zwischen Uralisch und Indogermanisch im nordwestindogermanischen Raum". – Alfred Bammesberger & Theo Vennemann (eds.): *Languages in Prehistoric Europe*. Heidelberg: Winter. Pp. 279-317.
- Kortlandt, Frederik H.H., 2003. "Initial Laryngeals in Anatolian".  
<http://www.kortlandt.nl/publications/art202e.pdf> [Also available in print: *Orpheus* 13/14: 9-12 (2003/2004).]
- Kroonen, Guus, 2012: "Reflections on the o/zero-ablaut in the Germanic iterative verbs" – Craig Melchert (ed.): *The Indo-European Verb*. Wiesbaden: Reichert. Pp. 191-200.
- Kuiper, Franciscus Bernardus Jacobus, 1962: "Atharvavedic *abhvā-*, n. "monster"". – *Lingua* 11: 225-230.
- Le Feuvre, Claire, 2004: "Note sur l'adjectif ἄργυρος, γύφος, ἀργύρεος, γύφρεος". – *Revue de philologie, de littérature et d'histoire anciennes* 2004/2, Tome LXXVIII: 257-264.
- Lehtisalo, T., 1936: *Über die primären Uralischen Ableitungssuffixe*. [= *Mémoires de la Société Finno-Ougrienne* 72]. Helsinki: Suomalais-Ugrilainen Seura.
- Lubotsky, Alexander, 2002: "The Indo-Iranian word for 'shank, shin'. – *Journal of the American Oriental Society* 122: 318-324.
- M&A = James P. Mallory & Douglas Q. Adams (eds.), 1997: *Encyclopedia of Indo-European Culture*. London / Chicago: Fitzroy Dearborn.
- McCone, Kim, 2005: "Mögliche nicht-indogermanische Elemente in den keltischen Sprachen und einige frühe Entlehnungen aus indogermanischen Nachbarsprachen". – Meiser & Hackstein 2005, pp. 395-435.
- Meier-Brügger, Michael, 1991: "Griechisch ἀφύη "Bratfischchen", ved. *ābhva-* "Unding", myk. Ortsname *a-phu-*". – *MSS* 52: 123-125.
- Meiser, Gerhard, 2003: *Veni Vidi Vici. Die Vorgeschichte des lateinischen Perfektsystems*. Munich.
- Meiser & Hackstein 2005 = Gerhard Meiser & Olav Hackstein (eds.): *Sprachkontakt und Sprachwandel. Akten der XI. Fachtagung der Idg. Gesellschaft*. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Melchert, Craig, 1991: "Death and the Hittite king". – Roger Pearson (ed.): *Perspectives on Indo-European language, culture, and religion. Studies in honor of Edgar C. Polomé* [= *Journal of Indo-European Studies monograph* 7]. Vol. 1, pp. 182-188.
- Melchert, Craig, 1994: *Anatolian Historical Phonology*. Leiden: Rodopi.
- Melchert, H. Craig, 2007: "Hittite *huwapp-*, *huppa(i)-* and *huppiya-*". – D. Groddek & M. Zorman (eds.): *Tabula Hethaeorum* [Fs. Košak]. Wiesbaden. Pp. 513-519.
- Michel, C. & M.-L. Nosch, 2010: *Textile Terminologies in the Ancient Near East and Mediterranean from the first to the third Millennium*. Oxford / Oakville.
- Neu, Erich, 1998: "Hittite *ši-mu-uš*". – *HS* 111, 1: 55-60.
- Neumann, Günther, 1988: "Beobachtungen an karischen Ortsnamen". – F. Imperati (ed.): *Studi di storia e di filologia anatolica dedicati a Giovanni Pugliese Carratelli*. Firenze.
- Oettinger, Norbert, 1979: *Die Stammbildung des hethitischen Verbums*. Nürnberg.
- Olsen, Birgit Anette, 1993: "Vedic and Laryngeals: *irajyati* and *iradhanta* – *badhnāti* and *ubhnāti*". – G. Meiser (ed.): *Indogermanica et Italica*. Innsbruck. Pp. 362-372.

- Olsen, Birgit Anette, 1994: "Armenian *dalowkn* 'jaundice' and the IE suffixes \*-g<sup>w</sup>on-, \*-gon- and \*-don-". – Jens Elmegård Rasmussen (ed.) in cooperation with Benedicte Nielsen: *In honorem Holger Pedersen. Kolloquium der Indogermanischen Gesellschaft vom 26. bis 28. März 1993 in Kopenhagen*. Wiesbaden: Dr. Ludwig Reichert Verlag, pp. 331-347.
- Olsen, Birgit Anette, 1999: *The Noun in Biblical Armenian* [Trends in Linguistics Studies and Monographs 119] Berlin / New York: Mouton de Gruyter.
- Olsen, Birgit Anette, 2000: "Verb or Noun? On the Origin of the Third Person in IE". – Martin E. Huld & al.: *Proceedings of the Twelfth Annual UCLA Indo-European Conference* [= *JIES Monograph Series* 40]. Washington D.C.: Institute for the Study of Man. Pp. 65-79.
- Olsen, Birgit Anette, 2003: "Another Account of the Latin Adjectives in *-idus*". *HS* 116 (2): 234-275.
- Osthoff, Hermann, 1877: "Etymologisches, Lautliches und Grammatisches". – *ZVS* 23: 84-94 [90-94: "Die Gotischen Adverbia auf *-o* und *-ba*"].
- Prellwitz, W., 1896: "Studien zur indogermanischen etymologie und wortbildung". *Beitr. zur Kunde der indogerm. Sprachen* 22: 76-114.
- Preveden, Francis R., 1930: "Church Slavonic *korabъ* and Greek *κάραβος*". – *Language* 6 (4): 279-296.
- Pronk, Tijmen, 2009: "Sanskrit (v)*r̥ṣa-bhā-*, Greek ἄρσην, ἔρσιν: the spraying bull of Indo-European?". – *HS* 122: 170-181.
- Puhvel, Jaan, 1984: *Hittite Etymological Dictionary*, I: *Words beginning with A* [= Trends in Linguistics; Documentation, 1]. Berlin: Mouton de Gruyter.
- Rasmussen, Jens Elmegård, 1989: *Studien zur Morphophonemik der indogermanischen Grundsprache*. Innsbruck: IBS.
- Rasmussen, Jens Elmegård, 2009: "Internal Reconstruction Applied to Indo-European: Where do we stand?". – Jens Elmegård Rasmussen & Thomas Olander (eds.): *Internal Reconstruction in Indo-European*. Copenhagen: Museum Tusculanum Press.
- Sadovski, Velizar, 2012: "Roots deep in heaven: Indo-Iranian ritual concepts in a cross-cultural perspective". Paper given at *Etymology and the European Lexicon*, the xth Fachtagung of the Indogermanische Gesellschaft, Copenhagen xx-xx 2012.
- Schindler, Jochem, 1967: "Idg. \**dheudh-* in Farbbezeichnungen". – *ZVS* 81, 1/2, 68-71.
- Schindler, Jochem, 1978: "Hittite *šalpa-*". – *Die Sprache* 24: 45.
- Schmalstieg, William R., 1996: Review of Ambrazas 1993. – *Lituanus* 42: 1.
- Shevoroshkin, Vitaly, 1979: "On the Hittite-Luwian Numerals". *JIES* 7 (3-4): 176-198.
- Sturtevant, Edgar Howard, 1911: "Studies in Greek Noun-formation: Labial Terminations". – *Classical Philology* 6, 2, pp. 197-215.
- Simon, Zsolt, 2008: Review of Adiego. – *Acta Antiqua Academiae Scientiarum Hungaricae* 48: 457-463.
- Stachowski, Kamil, 2008: *Names of Cereals in the Turkic languages* [= *Studia Turcologica Cracoviensia* 11]. Kraków: Księgarnia Akademicka.
- Watkins, Calvert, 1969: "A Latin-Hittite Etymology". – *Language* 45, 2: 235-242.
- Watkins, Calvert, 1995: *How to kill a dragon*. Oxford.

## On the Precursors of Celtic and Germanic<sup>1</sup>

### Abstract

A “Celto-Germanicism” may be defined as a lexeme shared by Celtic and Germanic only and appearing to be older than the emergence of Proto-Celtic and Proto-Germanic. Around 100 such items can be found, some of which are IE archaisms, while others look like morphological or semantic innovations. There is also a group of isolated lexemes which appears to have been borrowed from the same third source. Four fifths fall into two semantic spheres: 1) religion and healing, and 2) warfare and equestrian terminology. Most remarkable is the occurrence of as many as ten common words for ‘wound, injury’. This situation must reflect close contacts between speakers of the Indo-European dialects that later evolved into Proto-Celtic and Proto-Germanic respectively. We may tentatively fix this cultural unity in time and space in Eastern Central Europe around 2000 BCE, when the pre-Celtic Ūnétique culture bordered late varieties of the Corded Ware culture. Some shared loanwords can be traced back to Proto-Fennic, suggesting a continuum stretching further to the North. There are even indications that Proto-Fennic may have been in direct contact with Pre-Proto-Celtic, not always with Pre-Proto-Germanic as the provider. Words of possible Fennic origin include PCelt. *\*lubī-* ‘wort’ ~ PGmc. *\*lubja-* ‘poisoning or healing plant’, PCelt. *\*sanesto-* ‘secret advice’, PCelt. *\*magos* ‘plain, open field’, Nlr. *lón* ‘lunch’, PCelt. *\*klamo-* ‘disease’, and PGmc. *\*haljō-* ‘abode of the dead’.

---

<sup>1</sup> The present article was published as Hyllested 2010. Apart from this footnote (including the reference just mentioned), the abstract, the comments on Zair’s (2012) review and the inclusion of a new item no. 5) *\*uitelo-* (shifting each subsequent footnote by one), the two articles are identical.



### 1 Loanword or heritage ?

While Germanic has quite a few Celtic loanwords (see, e.g., de Vries 1960, Birkhan 1970, Mees 1998, Rübkeil 2002, Schumacher 2007), the share of older Germanic material in Celtic is comparatively small (Lane 1933: 264, and Schumacher 2007: 174-176). However, Celtic and Germanic also share lexical material exclusive to these branches that can be independently traced back to an identical reconstructed protoform. Therefore, it is often hard to determine whether a given Celto-Germanicism is inherited from PIE or borrowed from one branch to the other at a later age. Karsten (1927: 126) wrote on PGmc. *\*arbja-* vs. PCelt. *\*orbios* ‘heir’ and PGmc. *\*aiþa-* vs. PCelt. *\*oito-* ‘oath’: “kan likaså vara antingen urbesläktat med eller lån [might just as well be inherited as borrowed]”. Krahe (1954: 142) used the same lexeme as an example: “Die Hauptmasse des gemeinsamen nur keltisch-germanischen Wortschatzes reicht – ohne daß vom rein linguistischen Standpunkt Anhaltspunkte für eine Entlehnung aus der einen in die andere Sprache gegeben werden könnten – bis vor die Periode der Lautverschiebung zurück (Typus got. *aiþs* – air. *ōeth* usw.)”. Olsen (1988: 13) writes on PGmc. *\*gīslō-* ‘hostage’ vs. PCelt. *\*geistlo-* ‘id.’: “It is not certain whether the Gmc. examples are inherited or Celtic loanwords”. Casaretto (2004: 318, fn. 1051) on PGmc. *\*rū-nō-* ‘secret’ vs. PCelt. *\*rū-nā-* ‘id.’: “Ob diese Parallelität Lehnbeziehungen oder ein gemeinsames Erbe reflektiert, ist unsicher”. Ringe (2006: 296) states: “There are also quite a few words shared only by Celtic and Germanic, which might or might not be loanwords ...”. Matasović (2009: 227) on Proto-Celtic *\*krumbo-* ‘round, curved’: “Germ. *krumm*, OE *crumb* ‘round’ point to PGerm. *\*krumba-*, which was borrowed either from Celtic, or from the same non-IE source as the Celtic words”. Polomé (1983: 284) summed up the problem complex, listing four possible origins of a Celto-Germanicism: “a) the terms represented either a common regional innovation in a marginal area of the Indo-European territory or the localized survival of an archaic term lost elsewhere throughout the Indo-European Linguistic area; b) the terms have both been taken over from a same third source – be it a Pre-Indo-European (‘substrate’) language or less well-documented Indo-European language in their vicinity; c) the Celtic term was borrowed by Germanic; d) the Germanic term was borrowed by Celtic.” Lane (1933) and Elston (1934) excluded borrowing, i.e. possibility c) and d), when-

ver it could not be proved directly.<sup>2</sup> In the following, I will use the term “Celto-Germanicism” for items believed to be older than the emergence of Proto-Celtic and Proto-Germanic, but shared by these two branches only, i.e. Polomé’s categories a) and b).

## 2 Semantic spheres

Scholars already noted long ago that such Celto-Germanicisms pertain to certain semantic spheres. Thus, Lane (1933) suggested the following headings:

- a) Political and legal vocabulary
- b) Warfare
- c) Cultural and technical vocabulary, dwelling
- d) Nature, earth, land, plant and animal life
- e) Motion, locomotion, transportation
- f) The body and bodily functions
- g) Mental and emotional activity, vocal utterance
- h) Sense perception
- i) Family
- j) Religion, superstition
- k) Miscellaneous

and Krahe (1959: 139-141):

- a) Religion und Geistiges Leben
- b) Pferdezücht und Reiten
- c) Siedlung, Hausbau
- d) Landschaft, Natur
- e) Metalle
- f) Sonstiges

Elston (1934) and Campanile (1970) had still other divisions. On one hand, it is interesting to observe how an overrepresentation of shared vocabulary in certain semantic fields hints at the character of the rela-

---

<sup>2</sup> Schmidt (1984, 1986, 1987, 1991) proposed a five-strate model: *stratum 1*, whose Celtic origin is proved by their form; *stratum 2*, Celto-Germanic isoglosses with the same semantic shift; *stratum 3*, Celto-Germanic isoglosses without the same semantic shift; *stratum 4*, a group with special problems in the semantic field of craftsmanship; and *stratum 5*, name-doubles.

tionship in question. On the other hand, it seems as if Lane's and Krahe's lists cover most parts of the lexicon. If the lexical commonalities could be combined with shared innovations and archaisms in the phonological or grammatical system, it would be the obvious thing to hypothesize that Celtic and Germanic formed a subgroup within the Indo-European family. This seems not to be the case<sup>3</sup>. A closer scrutiny of the material indeed reveals a much less blurred and much more unambiguous picture of the character of the earliest Celtic-Germanic relations.

### 3 The material

Lists of Celto-Germanicisms have been compiled by Lane (1933), Elston (1934), Krahe (1956), de Vries (1960), Chemodanov (1962), Campanile (1969, 1970), and Polomé (1983). What follows is a revised and updated synthesis of their material with the addition of new items. Celto-Germanicisms include a) lexemes with a specialized meaning or use shared by Celtic and Germanic, b) formations particular to Celtic and Germanic (although formed from well-known PIE elements) and c) lexemes whose very roots or root variants are unknown outside Celtic and Germanic. The items have been grouped according to meaning (see section 4 below), but within these semantic groupings they appear in a more or less random order<sup>4</sup>.

<sup>3</sup> The features mentioned by Schmidt (1991: 146-147) are either too weak or too common to count as obviously shared innovations.

<sup>4</sup> Space does not allow a word-to-word treatment of items that I have refused to include as true Celto-Germanicisms. A few examples may serve as prototypes: Craig Melchert (p.c.) kindly points out to me that *\*tegu-* 'thick' in OIr *tiug*, W *tew* ~ OE *þiċce*, OHG *dicki*, ON *þykk* 'thick' vs. Lith. *tānkus* id. < *\*tenk-* (-*u-/to*)- is probably also attested in Hittite *tagu-* 'thick, swollen' < *\*togu-* (Neu 1995); *\*luH-s* 'louse' > W *llau* 'lice' (< collective *\*luuā* < *\*luH-eh<sub>2</sub>*) Corn. *low*, Bret. *laou* id., OW *leu-esicc* 'louse-eaten' ~ ON *lús*, OE, OHG *lūs* 'louse' vs. Toch. B *luwo* A *lu* pl. *lwā* 'animal' probably also forms the basis of Lith. *liūlė* 'louse' (where *-lė* is diminutive, cf. *brolė* 'brother'); PCelt. *\*korkio-* 'oats' (believed by Matasović 2009 to be of a common substratum origin) corresponds to Shughni *sip(i)yak* 'a kind of millet', *sepyak* 'grain of wheat' according to Stalmaszyk and Witczak (1991-1992); Rasmussen (1998) regards PGmc. *\*landa-* '(open) land' as a borrowing from Celtic proper. Despite the intriguing similarity, OIr. *nasc* 'ring; clasp; bond, tie' (~ *nascim* 'to bind') is most likely unrelated to OHG *nusca*, OS *nuscia* 'clasp, buckle' which is rather a Balto-Fennic borrowing, cf. Fi. *nuska*, *nurkka* 'corner, nook', especially since another word for 'buckle', ON *sylgja*, is already known to originate from Balto-Fennic (for the semantics, compare the

## Category 1

## A. Unique meaning

- 1) PCelt. *\*soito-* ‘sorcery’ > MW *hut*, Bret. *hud* ‘magic’ ~ PGmc. *\*saiða-* > ON *seiðr* ‘magic; spell, charm, enchantment’, *siða* ‘work a charm through *seiðr*’. Probably identical to PIE *\*soǵ-to-* ‘string, rope’, derived from *\*seh<sub>2</sub>ǵ-* ‘to bind’<sup>5</sup>, cf. Lith. *saĩtas* ‘sign, soothsaying, soothsayer, talisman’, but in Baltic also still ‘string, necklace etc.’, cf. Lith. *siẽtas*, Latv. *saĩte* id.
- 2) PCelt. *\*oitos* ‘oath’ > OIr *óeth*, MW *an-udon* ‘perjury’ ~ PGmc. *\*aiþaz* ‘oath’ > OE *āþ*, OHG *eid*, ON *eiðr*, Goth. *aiþs* ‘oath’, OS *mēn-ēþ* ‘perjury’; vs. Gk. *oĩros* ‘faith’, all from PIE *\*h<sub>1</sub>óǵ-to-s* ‘walk(ing)’, derived from *\*h<sub>1</sub>ei-* ‘to go’, cf. ON *ganga eið* ‘take the oath’ (see also Schumacher 2007: 176-177;)<sup>6</sup>.
- 3) PIE *\*kor-ǵo-no-* in the epithet of a god: OBrit. tribal name *Coriono-tōtae* ‘people of the army-lord (a god, probably Lugus)’ ~ PGmc. *\*harjanaz* > ON *Herjann* ‘lord of the army’, epithet of Odin; vs. Gk. *koĩpavos* ‘ruler, commander’ < *\*koĩpovos* (Meid 1991: 48-49)
- 4) A personalized form meaning ‘ghost’ of PIE *\*d<sup>h</sup>rrouǵhós*: OIr. *air-drech* ‘phantom’, MIr. *aur-fraich* ‘ghost’ ~ ON *draugr* ‘ghost’ vs. Ved. *drógha-* ‘deceiving’, Av. *draoga-* ‘lie’ (Mallory & Adams)<sup>7</sup>.
- 5) PCelt. *\*uitelo-* > MIr. *fiothal* ‘dwarf, hag, goblin; anything stunted’ ~ PGmc. *\*wīþila-* > OE *wīdl* ‘impurity’, OHG *widil*, *widillo* ‘homosexual, hermaphrodite, effeminate male’; both personified, vs. Lat. *vītiligō* ‘psoriasis, skin affliction’, deri-

---

double meaning of Da. *krog* ‘nook; hook’ and Eng. *nook*, a Scandinavian loan with the original meaning ‘clasp, hook’, ON *hnokki*).

<sup>5</sup> See Rasmussen 1989: 59-60.

<sup>6</sup> Nicholas Zair (p.c.) points out to me that since no derivative of *ganga* in itself means ‘oath’, *ganga eið* does not in itself suggest that *\*Hóǵ-to-* is derived from ‘to go’, and a connection with the root of Hitt. *hā(i)-* ‘believe’ must also be considered (Puhvel 1991: 10). However, this does not affect its status as a Celto-Germanicism.

<sup>7</sup> OIr *-drech* may also be identical to *drech* ‘vision’ < PCelt. *\*drikā*, derived from *\*dṛǵ-* ‘to see’.

ved from *\*vītis* ‘affliction’ (Kerckhof 2012)<sup>8</sup>. Derivatives from *\*ueih<sub>1</sub>*- ‘to twist’, cf. also Lat. *vitium* ‘defect, fault’.

#### B. Unique morphology

- 6) PCelt. *\*rūna-* > OIr. *rún* ‘secret’, MW *rhin* ‘spell, enchantment’ ~ PGmc. *\*rū-nō-* ‘secret’ > OE *rún*, OS *rūna* ‘whisper; secret; a rune’, ON *rún* ‘rune; secret’<sup>9</sup>.
- 7) Gaul. (Chamalières) *ande-dion uēdiū-mi* ‘I praise a god’ ~ Goth. *in-weitan gub* ‘to praise God’ < *\*uēid-* ‘see’ where *ande-* semantically corresponds to *in-* (de Bernardo Stempel 2001).<sup>10</sup>
- 8) PCelt. *\*nem-eto-* ‘sacred grove, sanctuary’ > Gaul. *νεμήτον* (Vaison), OIr. *nemed* ‘sanctuary’ ~ PGmc. *\*nemiþa-* > OS, OLFr. *nimidas* ‘sacred grove’ vs. Lat. *nemus* ‘sacred grove’, Gk. *νέμος* ‘wooded pasture, glade’.
- 9) PCelt. *\*daunā* > MIr. *dúan* ‘poem’ ~ ON *tafn* ‘sacrificial animal’ < *\*dap-no-* vs. Lat. *daps* ‘sacrificial meal’ < PIE *\*dap-* (Watkins 1970)<sup>11</sup>.
- 10) PCelt. *\*uiro-k(w)ū-*, gen. *-kunos* ‘werewolf’ > Celtib. (Botorrita III) **uiroku**, OIr. *Ferchú*, OW *Gurci*, OBret. *Gurki* ‘(name of a) werewolf’ (McCone 1987; McCone 2005: 401) ~ OE *wer(e)wulf*, Dan. *varulv*, Fr. *loup-garou* id. where *-garou* < Franconian *\*war-ulf-*.
- 11) PCelt. *\*nerto-* > OIr. *nert* ‘strength, force’, OW, MW *nerth*, Bret. *nerzh*, Gaul. PN *Nerto-maros* ~ Gmc. *\*ner-þu-* > goddess *Nerthus* ‘terra mater’, ON *Njǫrðr*, father of Freyr. Derivatives from PIE *\*h<sub>2</sub>ner-* ‘man; strong’ (Meid 1991: 15).

<sup>8</sup> The connection to Skt. *vetāla-* ‘demon’ (Lehmann 1907) is uncertain. If it is in fact related, the Celto-Germanic character of the present item should perhaps rather defined as a combination of morphology (found also in Latin) and semantics (found also in Indic).

<sup>9</sup> Rasmussen’s (1986: 1, 310) judgment that “the exact correspondence between Celtic and Germanic probably reflects an ancient borrowing in one direction or the other” is based on an isolated view of this lexeme. Contra Rasmussen’s connection with some Greek material, see Vine (2002: 206ff.).

<sup>10</sup> As she notes, Goth. *inweitan* takes the accusative while the Greek original takes the dative, i.e. chances are that this is not a Greek calque.

<sup>11</sup> The Celto-Germanic morphology also differs from Hitt. *tappala-* ‘person responsible for court cooking’, if it in fact belongs to the same root.

- 12) PCelt. \**ab-anko-* ‘water creature’ > OIr., MIr. *abac* ‘dwarflike creature associated with water’, W *afanc* ‘beaver’ ~ PGmc. \**abō(n)* ‘monkey etc.’ (Schrijver 2004).

### C. Isolated lexemes

- 13) PCelt. \**uāti-* > OIr. *fáith* ‘prophet’, Gaul. *ouáteis* (pl., Strabo) and \**uātu-* ‘shamanic wisdom’ > *fáth* ‘prophecy’, MW *gwawd* ‘ode’ ~ PGmc. \**wōð-* > ON *óðr* ‘poetry; furious’, Goth. *wōþs* ‘furious’, ON *Óðinn*, OE *Wōden*, OHG *Wuotan* ‘Odin’ (Meid 1991: 25-26; Watkins 1995: 118).<sup>12</sup>
- 14) PCelt. \**rīmā-* > OIr. *rím*, W *rhif* ‘number’ ~ PGmc. \**rīma-* > OE *rím* ‘number’, ON *rím* ‘computation’, OHG *rīm* ‘account, series, number’.
- 15) PCelt. \**sketlo-* > OIr. *scél* ‘tale’, W *chwedl* ‘saying, fable’ ~ PGmc. \**skapla-* > ON *skáld* ‘poet’.<sup>13</sup>
- 16) PCelt. \**gaisto-* > OIr. *gáes* ‘speculation’, cf. *gáeth* ‘insanity; wind’ ~ PGmc. \**gaista-* ‘(supernatural) spirit’ > OHG *geist*, OS *gēst*, OS *gāst* (Meid 1965).
- 17) PCelt. \**klamo-* > OIr. *clam*, W *claff* ‘grave’ ~ PGmc. \**skalmō* ‘plague, (cow’s) disease; evil spirit, crook’. Perhaps both from PIE \**skolm-eh<sub>2</sub>* ‘disease, evil spirit’, but the Proto-Celtic vocalism is not entirely clear; syllabic \*-l- preceding \*-m would normally yield \*-li-.<sup>14</sup>
- 18) PCelt. \**skāx-slo-* ‘demon, supernatural being’ > OIr. *scál* ‘phantom’, MW *yscawl* ‘young hero, warrior’ ~ PGmc. *skōh-sla-* > Got. *skohsl* ‘evil spirit, demon’; both from \**skōk-slo-*.
- 19) PCelt. \**buko-* > MIr. *boccánach* ‘goblin’, W *bwg* ‘ghost, hobgoblin’, *bwgan* ‘bogey, ghost’, *bwgwn* ‘fright’ ~ Fris. *bökk*, Swab. *bockelman*, NE *bogle*, *bogey*.

<sup>12</sup> I assume Lat. *vātēs* ‘prophet, seer’ to be a loan from Celtic.

<sup>13</sup> It is no longer necessary to reconstruct a labiovelar for this word to account for -w- in Welsh; cf. Schrijver 1992 and Jørgensen 2010. Zair (2012: 80) in his review objects that a connection with Lat. *insece* ‘say’ “still seems very plausible”. I do not quite understand this message, since I indeed do follow Schrijver and Jørgensen in leaving out the labiovelar, paving the way for both forms to match the Latin material. Even if Lat. *insece* is related, I will maintain that the nominal formation with \*-lo- (and its meaning) constitutes a Celto-Germanicism, albeit belonging to category II.

<sup>14</sup> If Alb. *helm* ‘poison’ belongs here, the Celto-Germanic connection is less clear.

## Category II

## A. Unique morphology (and meaning)

- 20) PCelt. *\*kol-ino-* > Ir. *cuilenn*, W *celyn* ‘holly’ ~ PGmc. *\*hul-isa-* > OE *holezn*, OHG *hulis*, OFr. *\*huls* > Fr. *houx* ‘holly’; vs. OCS *\*klasŭ* ‘ear of grain’, Toch. B *klese* ‘barley meal’, Alb. *kallí* ‘straw, chaff’, Skt. *kaṭamba-* ‘arrow’, all from PIE *\*kel-* ‘sharp, prickly’. According to Pliny, the plant was a popular house adornment among Celtic and Germanic peoples. In Germany and Austria, holly is traditionally placed in stables to protect horses from evil spirits<sup>15</sup>.
- 21) *\*k̥uond<sup>h</sup>-neh<sub>2</sub>* ‘Angelica’ > Ir. *cuinneóg* ‘wild angelica, *Angelica silvestris*’ ~ PGmc. *\*hwannō* > ON *hvōnn* ‘holy ghost, *Angelica archangelica*’ vs. *\*k̥uend<sup>h</sup>-ro-* with other meanings in Lith. *švėndras* ‘reed, reed-mace’; Lat. *combrētum* ‘a kind of rush’. Angelica is an old medicinal herb and was used against evil spirits (Birkhan 1999).

## B. Isolated lexemes

- 22) PCelt. *\*lubī* or *\*lubā* > OIr *luib* ‘wort, plant’ ~ OE *lybb*, OHG *luppi* ‘magic remedy; strong plant-juice; poison; magic’, ON *lýf* ‘healing plant’, Goth. *lubja-leisei* ‘magic; poisoning’. Perhaps also in ON *epli ellilyfs* ‘old-age medicine’ > *epli ellifu* ‘eleven apples’ (in the Eddic lay *Skírnismál*; see Polomé 1994: 142–143 on the similar role of apples in Germanic and Celtic mythology).

## Category III

## A. Unique meaning

<sup>15</sup> The holly’s connection to both horses and evil spirits may be due to the near-homonymy of PGmc. *\*marha-* m. ‘horse’, *\*marhjō-* ‘mare’ and *\*marō-* f. ‘female incubus’, let alone their complete homonymy in Scandinavian—cf. the ambiguous names of the holly, Dan. *maretorn*, *mareved*, *maretidse*, *marelok*, Nw. *marekvist*, Sw. *markvist*, *marlock*, *martova*, Icel. *marhrísla*, MLG *marvlechte*, *marlock*, *mahrzopf*. Other Germanic names refer to the spirits only: Nw. *huldretime*, NHG *Schrattelbaum*, *Hexenbesen*, Eng. dial. *witch’s besom*.

- 23) PIE *\*stér-b<sup>h</sup>eh<sub>2</sub>* ‘stiffness’ in the specialized meaning ‘death; plague’: OIr. *us-sarb* ‘death’<sup>16</sup> ~ OHG *sterbo*, OE *steorfa* ‘plague’ vs. Gk. *στέρφος* ‘animal skin, leather’, all from *\*ster-* ‘(be) stiff’.

#### B. Unique morphology

- 24) *\*(H)órbh-īo-* m. ‘leavings’ and *\*(H)órbh-īo-m* n. ‘inheritance’ > OIr. *orbae* ‘inheritance’, Gaul. *Orbio-* ‘id.’, OIr. *orb(b)e*, *orpe* ‘heir; inheritance’ ~ Goth. *arbja* ‘heir’, *arbi* ‘inheritance’, OHG *erbi*, OE *ierfe* ‘id.’ vs. *\*(H)orbh-o-* ‘orphaned’ > Lat. *orbus* ‘deprived’, Gk. *ὀρφανός* ‘orphaned’, Arm. *orb* ‘orphan’, Skt. *árbha-* ‘small; weak; child’ (McCone 1999).
- 25) PIE *\*krejH-* in *\*kriH-no-* > PCelt. *\*krīno-* > OIr. *crín* ‘enfeebled by old age, decrepit; withered’, OW *crin* ~ *\*kroiH-uo-m* > PGmc. *\*hraiwa-* n. > Goth. *\*hraiw* in *hraiwa-dubo* ‘turtle dove’, ON *hræ* ‘dead body’, OE *hræw* ‘id.’, OHG *hrēo* ‘dead body; grave; funeral; death’ vs. the unextended root *\*kerh<sub>2</sub>-* ‘to break’ (Casaretto 2004: 164).
- 26) PCelt. *\*uer-t-* > OIr. *fertae* (< *\*-iā*) ‘burial mound’, W *gwerthyr* ‘fort’ (< *\*-ero-*) *gweryd* (< *\*-eto-*) ‘earth, soil; grave’ ~ OE *weorð* ‘yard’, *weard* ‘guarding’, ON *varða*, *varði* ‘milestone’, *vǫrðr* ‘warden, watchman, defender; guardian spirit, house spirit, soul of the dead’.

#### C. Isolated lexemes

- 27) PCelt. *\*doueno-* > OIr. pl. *dóini*, *doíni* ‘men’, poetic sg. *doín*, *doén* ‘man’<sup>17</sup> ~ PGmc. *\*dewena-* > Goth. *diwans* ‘mortal’, cf. the verb OHG *touwen*, OS *dōian* ON *deyja* ‘die’.
- 28) PCelt. *\*krito-* > OIr. *crith* ‘trembling; fever’, *crith-galar* ‘illness with fever’, W *cryd* ‘fear’ ~ PGmc. *\*hriða-* > OE *hrið* m. ‘fever’, Nw. *ri* ‘sudden illness; short period; hard weather’ (Bjorvand & Lindeman 2000: 724).

<sup>16</sup> OIr. *-rb-* in *us-sarb* may be from *\*-ry-* instead, cf. *marb* ‘dead’ < *\*mr̥-uo-s*.

<sup>17</sup> Historically a suppletive paradigm with the sg. *duine* from PCelt. *\*gdonio-* ‘earthling’ corresponding to Ved. *kṣámya-* ‘earthly, mortal’, cf. Gaul. TEVO-XTONION (Vercelli) ‘of god and men’. Even if Latin *fūnus* ‘funeral procession’ is related, the item still constitutes a Celto-Germanicism in terms of semantics and word-formation (cf. Rasmussen 1988:92–3).



- 29) PCelt. *\*trusko-* > OIr. *trosk* 'leprous; leper', W *trwsgl* 'rude; clumsy', Bret. *trousk* 'polyps' ~ Goth. *þruts-fill*, OE *ðrúst-fell* 'leprosy'.

#### Category IV

##### A. Unique meaning

- 30) PCelt. *\*gaiso-* 'spear' > OIr. *gái*, Gallo-Gk. *γαῖσον*, Gallo-Lat. *gaesum*; OIr. *fo-gae*, W *gw-aew* 'javelin' ~ PGmc. *\*gaizo-* > OE *gār*, OHG *gēr*, ON *geirr* 'dart, spear'; from PIE *\*g<sup>h</sup>áisos*, cf. Gk. *χαῖος* 'shepherd's crook', Skt. *heṣas-* 'weapon'.
- 31) PCelt. *\*prijo-* > MW *ryd*, OCorn. *rid* 'free' ~ PGmc. *\*frija-* > OE *frēo*, OHG *frī*, Goth. *freis* 'free' vs. Ved. *priyá-* 'dear' (Schumacher 2007: 177).
- 32) PIE *\*kel-* 'strike' used in words for 'battle': PCelt. *\*kellāko-* > Mlr. *cellach* 'contention, strife' ~ OE *hild* 'war, battle', OHG *hiltia*, ON *hildir* 'battle' (corresponding to the ethnonym *Celtae*).
- 33) PCelt. *\*trex-so-* > OIr. *tress* 'battle' and *\*trex-s-no-* > OIr. *trén* 'brave, strong', comp. sup. *tressa*, *tressam* ~ PGmc. *\*þrak-ja-* > OE *þrece* 'force, oppression', OS *wāpan-threki* 'ability with arms', ON *þrekr* 'strength, bravery'.
- 34) PCelt. *\*kagro-* 'enclosure, fort' > W *caer*, Mbret. *ker*; and *\*kagio-* 'pen, enclosure' > MW *cae* 'fence', OBret. *caiou* pl. 'fortification, bulwark' ~ PGmc. *\*hagan-* 'enclosure, fence' > ON *hagi* 'pasture with a fence', OE *haga*, OHG *hac* 'hedge' and *\*hagjō-* > OE *hecg* 'hedge'.

##### B. Unique morphology

- 35) PCelt. *\*drungo-* > Ir. *drong* 'troop', MW *dronn* 'multitude', Gallo-Lat. (Vegetius) *drungos* 'groups of enemies' ~ PGmc. *\*drūhta-* > OE *dryht* 'companion', OHG *truht* 'troop' ON *drótt* 'company, following', Goth. *driugan* 'to serve as a soldier' vs. OCS *drugŭ* 'friend, other' *draŭgas* 'friend'.
- 36) PCelt. *\*uik-* 'fight' > Ir. *fichim* 'fight', *fecht* 'military expedition', OW *guith* 'front' ~ Goth. *weihan*, OE, OHG *wīhan*

- ‘fight’, ON *vega* ‘kill, fight’ vs. Lat. *vincō* ‘conquer’, Lith. *veĩkti* ‘make, work’.<sup>18</sup>
- 37) PCelt. \**kauno-* ‘harbor’ > MIr. *cúan* ~ PGmc. \**haf-na-*; originally ‘enclosure, shelter (for vessels)’.
- 38) PCelt. \**baduo-* ~ \**boduo-* > OIr *Badb* ‘name of the slaughter goddess’,<sup>19</sup> Early Ir. *badb* ‘crow, demon, witch’, W *bod* ‘kite’, NIr. *badhbh* ‘vulture; hoodie crow; fairy; scold’, Gaulish deity *Catu-bodua* ~ PGmc. \**baðu-* ~ \**baðwa-* > OHG *Batu-* ‘slaughter; battle’ (in names), OE *beadu* ‘battle’, ON *bǫð*, *bǫðvar* ‘war’ vs. Lat. *fodiō*, Hitt. *padda-* ~ *padd-* ‘dig (the ground), bury’, Toch. A *pāt-* ‘plough’, OCS *bodǫ* ‘to stab’, *bedǫ* ‘to stick, to dig’.
- 39) PCelt. \**nītu-* > OIr. *níth* ‘battle, distress’, Gaul. PN *Nitio-broges*, *Nitio-genna* ~ PGmc. \**nīþa-* > Goth. *neip* ‘envy, enmity’, OE *nīþ*, OHG *nīd* ‘battle-rage, hate, envy’, ON *níð* ‘libel’.
- 40) PCelt. \**magu-* > Ogham *magu* ‘slave’, W *meu-dwy* ‘hermit (< ‘servant of God’)', MBret. *m(a)oues* ‘girl’ ~ Goth. *magus* ‘boy’, ON *mǫgr* ‘son; youth’, OE *magu* ‘child; son; man’.

#### C. Isolated lexemes

- 41) PCelt. \**nanti-* > OIr. *néit* ‘battle, combat’, *Néit* ‘god of battle, husband of the war-goddess Nemain or Badb’ ~ PGmc. \**nanþjana-* > OE *nēþan*, OHG *gi-nenden*, ON *nenna*, Goth. *ana-nanþjan* ‘to dare’.
- 42) PCelt. \**poiko-* > OIr *óech* ‘enemy’ ~ PGmc. \**faiha-* ~ \**faiga-* > OE *fáh*, *fág* ‘guilty; outlawed; hostile’, NE *foe*, OHG *fēhida* ‘hate, enmity’, Goth. *fáih* ‘deceit’.

<sup>18</sup> Brent Vine (p.c.) points out to me that while the nasal present in Lat. *vincō* appears (predictably) beside an old root aorist in perf. *vici*, in theory (despite LIV<sup>2</sup> 670-671) the Celtic and Germanic presents could also be derived from the old root aorist (e.g.: root aorist subjunctive → thematic present is well-attested). In that case, the Germanic and Celtic material might be closely related, morphologically, to the old aorist (as in Latin), and since the Latin semantics is quite similar to the one shared by Celtic and Germanic, only the development into a thematic present would then point to a Celto-Germanicism.

<sup>19</sup> Remarkably, *Badb* is the sister of *Macha*, married to *Nemed*, and of *Mor-rígain*; of these four names, the first three are all Celto-Germanicisms, while cognates of *Mor-* are also found in Slavic.

- 43) PCelt. \**slak-* ‘strike’ > MIr *slactha* ‘struck (ptc.)’, *slacc* ‘sword’, Gael. *slachdaim* ‘strikes with a hammer’ ~ PGmc. \**slahana-* > Goth., OHG *slahan*, ON *slá*, OE *slēan* ‘slay’.
- 44) W *llost* ‘spear’, Bret. *lost*, Ir. *loss* ‘end; tail’ ~ ON *ljóstr* ‘fish-spear’, Dan. *lyster* ‘eel-spear’, *ljósta* ‘strike’.
- 45) PCelt. \**mūg-* ‘conceal’ > OIr *for-múigthe*, *for-múichthai* ‘smothered, concealed’ ~ PGmc. \**mūk-* > OHG *mūhhen* ‘lie in ambush for’, NHG *Meuchler* ‘assassin’, ME *micHER* ‘thief’, Eng. dial. *mitch* ‘hide (oneself)’.
- 46) OIr *bágaid* ‘fight, boast’, *bág* ‘battle’, W *beio* ‘blame’, Gaul. *Bagaudae*, probably ‘the fighters’, name of Gallic peasants who rebelled under Diocletian ~ PGmc. \**bēg-* > OHG *bāgēn* ‘quarrel, fight’, ON *bægjast* ‘quarrel, strive’.<sup>20</sup>
- 47) PCelt. \**gwelti-* ‘madman, lunatic’ > MIr. *geilt* ‘panic-stricken fugitive from battle’, W *gwyllt* ‘wild, savage, mad’ ~ Goth. *wilpeis*, OE *wilde*, OHG *wildi*, ON *vildr* ‘wild’.
- 48) PCelt. \**ueidu-* ‘wild’ > OIr. *fiad* ‘wild animals’, *fian* ‘troop of young warriors’, MW *gwydd* ‘wild’, *gwyddel* ‘a Gael, Irishman’ ~ PGmc. \**wajðaz* > OE *wāð* ‘hunt’.
- 49) PCelt. \**boudi-* ‘victory’ > OIr. *búaid* ‘victory’, W *buddig* ‘victorious’ ~ ON *býta* ‘exchange, divide’, MLG *būte* ‘booty’<sup>21</sup>, all from \**b<sup>h</sup>óud<sup>h</sup>i-*.
- 50) PCelt. \**pleid-o-* ‘succeed’ > MW *llwyddaw* ~ PGmc. \**flitana-* > OE *flitan*, OHG *flīzan* ‘attempt, try hard’.
- 51) PCelt. \**geistlo-* ‘hostage’ > Ir. *gíall*, W *gwystl* ‘hostage’, Br. *gouestl* ‘vow; promise’, Gaul. PN *Con-geistlus* ~ PGmc. \**gīslo-* > OHG *gīsal*, NHG *Geisel*, OE *gīsel*, ON *gísl* ‘hostage’.
- 52) PCelt. \**dūno-* ‘fortification, rampart’ > Ir. *dún*, W *din*, Gaul. *-dūnum* in place-names ~ ON and OE *tún* ‘hedged or fenced lot, enclosure’; OHG *zūn* ‘enclosure, hedge’.<sup>22</sup>
- 53) Ir. *clíab* ‘\*shield (of wicker-work)’ > ‘basket; wicker frame of a boat; chest’ ~ ON *hlíf* ‘shield, protection’, OHG *līpen*, *līpen* ‘protect’, Goth. *hleibjan* ‘take the part of’.

<sup>20</sup> Even if Latv. *buōzties* ‘be annoyed’ belongs here (LIV<sup>2</sup> 68 \**b<sup>h</sup>eh<sub>1</sub>ġ<sup>h</sup>-*), Celtic and Germanic still share a common semantics.

<sup>21</sup> NE *booty* is a borrowing from Scandinavian.

<sup>22</sup> Even if these words are derived from a PIE root \**d<sup>h</sup>eu<sub>h</sub>h<sub>2</sub>-* ‘be finished, come full circle’ (Watkins 1991:453), the derivative and its meaning are specific to Celtic and Germanic.

- 54) NIr *tailm*, Bret. *talm* ‘sling’, W *telm* ‘snare, trap’ ~ ON *þjálmi* ‘sort of snare’. The ON consonantism seems to indicate common heritage.

### Category 5

#### A. Unique meaning

- 55) PCelt. *\* Brus-na-* > OIr. *bronnaim* ‘injure, damage’, *\* Brus-o-* > W *briw*, Corn. *brew* ‘wound’ ~ PGmc. *\* Brūs-* > OE *brȳsan* ‘bruise’, OHG *brōsma* ‘crumb’ vs. Lat. *frustum* ‘fragment’.
- 56) PCelt. *\* kaiko-* ‘having an eye defect’ > OIr. *cáech*, OCorn. *cuic* ‘one-eyed’, W *coeg-ddall* ‘half-blind’ ~ Goth. *háih* ‘one-eyed’; vs. Lat. *caecus* ‘blind’. The Celtic god Lug closes one eye in his magic ritual, while in Germanic mythology, Odin is one-eyed (Polomé 1994: 145).
- 57) PCelt. *\* knidā-* > OIr. *cned* ‘a wound’ ~ PGmc. *\* hnītana-* > ON *hníta* ‘wound to death’, OE, OS *hnītan* ‘thrust, stab’; vs. Gk. *κνίζω* ‘to scratch’. ‘servant of God’),

#### B. Unique morphology

- 58) PCelt. *\* aglo-* ‘wound, affliction’ > OIr. *áil* ‘insult’, MIr *\* álad* ‘wound’, MW *aeled* ‘pain; grief’ ~ PGmc. *\* agla-* > OE *egle* ‘disagreeable, loathsome’, Goth. *agls* ‘shameful’, *aglipa*, *aglo* ‘affliction’ vs. Av. *ayō* ‘bad, evil’, Skt. *aghá-* ‘bad’, *aghrā-* ‘evil, distress’, *aghalá-* ‘terrible’, all from PIE *\* agh-* or possibly *\* h<sub>2</sub>egh-*.
- 59) PCelt. *\* g<sup>w</sup>en-i-* ‘wound’ > OIr. *guin* ‘wound, injury’ ~ PGmc. *\* banjō-* > Goth. *banja* ‘strike, wound’, ON *ben*, OE *ben(n)* id., OS *beni-wunda* ‘wound’ vs. PGmc. *\* ban-an-* ‘murder’ in OE *bana*, Da. *bane-sår* ‘deadly wound’ < PIE *\* g<sup>h</sup>wen-* ‘to kill’. wound to death’.
- 60) PCelt. *\* koldo-* > OIr. *coll* ‘destruction’, W *ar-choll* ‘wound’ ~ PGmc. *\* halta-* > Goth. *halts*, OE *healt* ‘lame’.
- 61) PCelt. *\* kre(n)x-tu-* > OIr. *crécht* ‘wound’, W *creithen* ‘scar’, MBret. *creizenn* id. ~ PGmc. *\* skranh-a-* > ON *skrá* ‘scroll’.

#### C. Isolated lexemes

- 62) W *gwanu* ‘pierce, thrust, stab’, *ym-wan* ‘joust, tilt’, *ymwanwr* ‘combatant’ ~ PGmc. *\*wunda-* > Goth. *wunds*, OE *wund*, OHG *wunda*, ON *und* ‘wound’.<sup>23</sup>
- 63) PCelt. *\*snad-o-* > MIr *snaidid* ‘to cut; to scratch’, W *neddyf* ‘axe’ ~ OHG *snatta* ‘wound; scar; bruise’, ON *snata* ‘spear’.
- 64) PCelt. *\*saitro-* > OIr *saethar* ‘work, labour’ and *\*saitu-* > OIr *saeth* ‘trouble’, MW *hoed* ‘pain’ ~ PGmc. *\*sai-ra-* > OE *sārig* ‘sorry’, OHG *serō* ‘painfully; in a difficult way’, ON *sárr* ‘painful’, *sár* ‘wound’.<sup>24</sup>

#### Category 6

##### A. Unique meaning

- 65) PIE *\*kād-* in derivatives with the meaning ‘hatred’: PCelt. *\*kad-s-i-*, *\*kādo-* > MIr *caiss*, W *cawdd*, Bret. *cas* ‘hatred’ ~ PGmc. *\*hatiz-* > Goth. *hatis*, OE *hete*, OHG *haz* ‘hatred’, ON *hatr* ‘hatred; persecution’ vs. Av. *sādra-*, Gk. *κῆδος* ‘sorrow; pain; misery’, Osc. gen.sg. **cadeis** ‘hostility’ (cf. also Birkhan 1967, Rübekeil 2001).<sup>25</sup>

##### B. Unique morphology

- 66) A secondary thematic derivative *\*h<sub>2</sub>up-él-o-* ‘evil’ > OIr *fel* ‘evil’ ~ PGmc. *\*ubila-* ‘evil’ > Goth. *ubils*, OE *yfel*, OHG *ubil* vs. Hitt. *ḫuwapzi* ‘ill-treats, spoils’, Toch. A *umpar* ‘bad’, all from PIE *\*h<sub>2</sub>uēp-* ‘treat badly’ (cf. Cohen & Hyllested 2007:16).
- 67) PCelt. *\*kloino-* > OIr. *cloen* ‘crooked; unfair; evil’ ~ PGmc. *\*hlaina-* ‘hill’ > Goth. *hlain* ‘hill’, Nw. dial. *hlein* ‘steep slope’, both with *\*-no-* from PIE *\*klei(H)-* ‘to lean’, cf. Ved. *śráyati*, Lith. *šliėti* ‘id.’, *šlainùs* ‘slanting’.

<sup>23</sup> Zair (2012: 80) in his review of Hyllested 2010

<sup>24</sup> Even if Lat. *saevus* ‘wild, ferocious’ and Hittite *šāi-* ‘be sullen, angry’ (see on the latter Kloekhorst 2008:692–3) are related, the Celtic and Germanic items form a semantic entity.

<sup>25</sup> Zair in his review (2012: 80) asks if the Celtic/Germanic meaning ‘hatred’ is unique enough to be seen as a shared feature compared to ‘hostility’ in the Oscean form. The two meanings differ in an important way, namely that hatred refers to a feeling while hostility refers to a behavior.

C. Isolated lexemes

- 68) PCelt. \**loktu* > OIr. *locht* 'fault, blame; mistake' ~ PGmc. \**lahana-* > OE *lēan*, OHG *lahan*, ON *lá* 'to blame'.

Category 7

A. Unique meaning

- 69) PCelt. \**reid-o-* 'ride; riding; chariot' > Gaul. *rēda* 'travelling-carriage with four wheels', OIr. *ríadaim* 'ride (in vehicle)', Ir. *dé-riad* 'team of two horses', W *rhwyddau* 'facilitate, speed' ~ ON *riða*, OE *rīdan*, OHG *ritan* 'to ride'; ON *reið* 'riding; horse-riding band; wagon' vs. Latv. *raidīt* 'send quickly; hunt'.

B. Unique morphology

- 70) Compounds with \**h<sub>1</sub>ek<sub>u</sub>o-* 'horse' and \**reid<sup>h</sup>-*: Gaul. PN *Epo-rēdo-rīx* ~ OE *eo-red*, OS *eo-rid-folc* 'cavalry', ON PN *Jó-reiðr*.
- 71) i-stem adjectives meaning 'easy', 'ready' derived from \**reid<sup>h</sup>-* via 'driving' or 'ready to go': PCelt. \**reidi-* > OIr. *réid* 'simple, easy, flat', W *rhuiidd*, OBret. *ruet* 'easy, quick' ~ OHG *bi-reiti* 'ready', Eng. *ready*.
- 72) PCelt. \**axsilā* > W *echel*, MBret. *ahel* 'axis' ~ PGmc. \**ahsulaz* > ON *ǫxull* 'axis' vs. formations without \*-*lo-* in Lat. *axis*, Lith. *ašis* 'id.'.
- 73) PCelt. \**uegno-* > OIr. *fén*, W *gwain*, Gaul. *co-vinnus* 'wagon' ~ PGmc. \**wagna-* > ON *vagn*, OHG *wagan* 'wagon' vs. other formations in Skt. *vāhana-*, Lat. *vehiculum* 'id.'.
- 74) PIE \**sent-* 'to travel' in nominal formations meaning 'road; retinue'; PCelt. \**sentu-* 'path' > OIr. *sét*, MW *hynt* 'path', *epynt* 'horseroad' ~ PGmc. \**sinþa-* and \**gasinþja-* 'retinue' > OHG *Gisindi* 'war retinue'.

C. Isolated lexemes

- 75) PCelt. *\*marko-* ‘horse’ > MIr. *marc*, W *march*, Bret. *marc’h*, Gaul. (acc.sg.) *μάρκav* ‘horse’, *Marco-* in place-names ~ OHG *mar(a)h*, OE *mearh*, ON *marr* ‘id.’.
- 76) PCelt. *\*drux-to-* > MIr. *drochta* ‘tub, vessel’ ~ PGmc. *\*trugaz* > OE, ON *trog*, OHG *troc* ‘trough’.
- 77) PCelt. *\*kanx-s-ikā-* > W *caseg*, Bret. *kazeg* ‘mare’, not formally identical to PGmc. *\*hangista-* ~ *\*hanhista-* ‘horse, stallion etc.’ (Jørgensen 2006), but their similarity can hardly be coincidental in the light of other equestrian commonalities; cf. also that PCelt. *\*keng-o-* ‘to tread, step, walk’ is irregular in the first place.
- 78) PCelt. *\*mongo-* ‘mane’ > MIr. *mong*, W *mwng* ‘id.’ ~ ON *makki* ‘upper part of a horse’s neck’, Dan. *manke* ‘mane’; cf. also ON *møn*, OE *manu* ‘mane’.
- 79) PCelt. *\*doklo-* > OIr *dúal* ‘strand, lock (of hair)’ ~ PGmc. *\*tagla-* > ON *tagl*, Dan. *tavl* ‘hair of a horse’s tail’, OE *tægl* ‘tail’, Goth. *tagl* ‘a hair’.

#### Category 8

##### A. Unique meaning

- 80) PCelt. *\*rādī-* > OIr *rádim* ‘to say, to speak’, MW *ad-raud* ‘to tell’ ~ PGmc. *\*rōðiana-* > Goth. *rodjan*, ON *róða* ‘to speak’.

##### B. Unique morphology

- 81) PCelt. *\*bana-tlo-* > W *banadl*, MBr. *malazn* ‘broom’ ~ OE *bōnian* ‘polish’, OS *bōnēn* ‘scrub, polish’, both from < *\*b<sup>h</sup>eh<sub>2</sub>-n-* ~ *\*b<sup>h</sup>ə<sub>2</sub>n-* vs. Gk. *φαίvw* ‘shine’, Arm. *banam* ‘open, reveal’ (Olsen 1988: 26).
- 82) PCelt. *\*gablo-*, *\*gablā-* > OIr *gabul*, NIr *gabhal*, W *gafl* ‘fork’ ~ PGmc. *\*gabalō-* > OHG *gabala*, OE *geafof* ‘id.’.
- 83) PCelt. *\*plāro-* ‘floor’ > OIr *lár* ‘ground, surface; middle’, W *llawr* ‘floor’, Bret. *leur* id. ~ PGmc. *\*flōruz* > OIr *flórr* ‘floor of a cow stall’, OE *flōr* ‘floor’ vs. Lat. *plānus*.
- 84) PCelt. *\*sītā-* > MIr. *sithlad* ‘sieving’, W *hidl*, MBret. *sizl* ‘sieve’ ~ PGmc. *\*sēpla-* > ON *sáld* ‘sieve’, Fi. (< Gmc.) *siekla*, *seula* id. all from *\*seh<sub>1</sub>ī-tlo-*, *\*sih<sub>1</sub>-tlo-* vs. *seh<sub>1</sub>ī-to-*, *\*sih<sub>1</sub>-to* in Lith. *sėtas*, CS *sito*.

- 85) PCelt. \**φatimā-*, \**φatamī-* > Gael. *aitheamh*, W *edafedd* ‘yarn; thread’ ~ PGmc. \**faþmaz* > OHG *fadum*, OE *fæðm*, ON *faðmr* ‘spread arms, embrace; thread’ (Hamp 2008).<sup>26</sup>
- 86) PCelt. \**iexti-* > Ir *icht* ‘tribe’, W *ieith* ‘language; nation’, MBret. *yez* ‘language’ ~ PGmc. \**jehti-* > OHG *jiht* ‘utterance’ (cf. *jehan* ‘to speak’) vs. Lith. *juōkas*, Lat. *iocus* ‘jest’.
- 87) PCelt. \**rextus* > OIr *recht* ‘law, justice’, MW *kyf-reith* ‘id.’ ~ PGmc. \**rehtuz* > ON *réttr* ‘justice, law’ < \**h<sub>3</sub>reǵ-tu-* with a shared, unpredictable meaning (Schumacher 2007:177).
- 88) PCelt. \**roino-* > OIr *roen* ‘road; mountain range’, Bret. *run* ‘hill’ ~ PGmc. \**raina-* > ON *-rein* ‘strip of land’ (in compounds), OHG *rein* ‘ridge of earth as boundary mark’.

#### C, Isolated lexemes

- 89) \**sʷek-* > W *chweg*, Bret. *c’houek* ‘sweet, pleasant (of taste)’, W *chwaeth* ‘taste’ ~ OE *swecc*, *swæcc* ‘taste, (pleasant) smell’, OHG *swehhan* ‘to smell (bad)’.
- 90) \**sʷem-* > OIr *to-seinn* ‘hunts; follows’<sup>27</sup> ~ OHG, OE *swimman*, ON *svim(m)a* ‘to swim’, Goth. *swum(f)sl* ‘lake’ < \**swum-sla-* (Bjorvand and Lindeman 2000:893–5, but they reject the connection; Casaretto 2004:408).
- 91) \**sʷeng-* ‘to bend’ in PCelt. \**swengo-* ‘slender’ > MIr. *seng*, Gaul. PN *Singi-dūnum* ~ OE *swancor*, MHG, MLG *swanc* ‘slender’, Dan. *svang* ‘arch of foot’ vs. \**sʷeg-* and \**sʷenk-* in other formations and languages (IEW 1047).

<sup>26</sup> Hamp includes Alb. *pê*, pl. *penj* ‘thread’, but Celtic and Germanic still agree both on o-grade and semantics.

<sup>27</sup> If Zair in his review (2012: 80), following LIV 532-3, is right that *to-seinn* is related to Hitt. *sanahzi* ‘sought’ < \**senh<sub>2</sub>-*, this item is of course not a Celto-Germanicism. Note also Kroonen 2013 forthc.: “The verb has no good extra-Gm. Etymology. The connection with OIr. *seinnid* is extremely doubtful, both on the formal and semantic side”. In fact, eliminating this item would only strengthen the hypothesis presented here since it reduces the number of items belonging neither to category 8. However, the combination of the initial consonant cluster \**sʷ-*, shared with items 89 and 91, and the derivational suffix \*-*slo-* of PGmc. \**swumsla-*, shared with items 19 and 51, perhaps point to an origin in the Celto-Germanic stratum after all. Indeed, the PGmc. ablauting forms \**swammjan-*, with a causative, not denominative, meaning ‘make swim’, \**swamōn* ‘swim’ and \**sunda-* ‘sound’ (beside \**swumsla-*), indicates that it is fairly old.



- 92) PCelt. *\*grando-*, *\*grendo-* ‘beard’ > MIr. *grend*; MW *grann* ‘beard; chin; cheek’, Provençal *gren* ‘moustache’ (< Gaul.) ~ PGmc. *\*granō-* f. > OE *granu* ‘moustache’, OHG *grana* ‘hair of the beard’, ON *grøn*, Goth. *grano* ‘hair of the beard; spruce (needle)’.
- 93) PCelt. *\*lind-o/u-* ‘drinkable water’ (cf. Matasović 2009:240) > OIr *lind* ‘liquid’, W *llyn* (m/f) ‘drink’, (m) ‘lake’ ~ Icel. *lind* ‘spring, fountain’, MHG *lünde* ‘wave’.
- 94) PCelt. *\*glāuo-* > MW *glo* ‘charcoal’ and PCelt. *\*goulo-* (< dissimilated from *\*glou-lo-?*) > MIr. *gúal*<sup>28</sup> ~ PGmc. *\*kula-*, *\*kulan-* ‘charcoal’ > ON *kol* (pl.), OE *col*, OHG *kolo*.
- 95) PCelt. *\*druxtu-* (< *\*drup-tu-*) > OIr *drúcht* ‘dew, a drop’ ~ PGmc. *\*drupa-* > ON *dropi*, OE *dropa*, OHG *tropfo* ‘drop’.
- 96) PCelt. *\*kaito-* ‘wood’ > OW *coit* OCorn. *cuit*, MBret. *coat* ‘forest, wood’, Gaul. PN Καυτόβριξ, *Cēto-briga*, *Eto-cētum* ~ PGmc. *\*haipja-* > Goth. *haiþi* ‘field, heath’, NHG *heis-ter* ‘small tree or bush’, ON *heiðr* ‘heath, moor’.
- 97) *\*gan(dʰ)-no-* in MIr *gann* ‘vessel, jug, pitcher’ ~ PGmc. *\*kannō* f. > ON *kanna*, OE *canne*, OHG *channa* ‘can, jug’.

#### 4 Revision of semantic areas and their implications

Our revised list may be said to fall into the following categories:

- (1)-(19) cosmology, spirits, supernatural creatures
- (20)-(22) medicinal herbs or plants connected to popular beliefs
- (23)-(29) sickness and death
- (30)-(54) battle and warfare, fortifications, weaponry
- (55)-(64) words for ‘wound’, ‘injury’, ‘defect’
- (65)-(68) hostility
- (69)-(79) equestrian terminology
- (80)-(97) words belonging to other parts of the vocabulary

Or, in a boiled-down version (excluding (80)-(97)):

- (a) religion and healing ((1)-(29))
- (b) warfare and equestrian terminology ((30)-(79))

<sup>28</sup> Even if PCelt. *\*goulo-* and the Germanic forms are related to Skt. *jvālati* ‘burns’, Toch. B *śoliye* ‘hearth’, Celtic and Germanic still share a specialized meaning.

Quite a few of the lexemes in question can be placed in either category. Remarkably, as many as nine words for ‘wound’ turn out to be Celto-Germanicisms in one way or another.

Only 1/5 of the lexemes fall outside the two main categories, and, with a couple of exceptions, even these are typical culture-words. Such a distribution militates against the possibility of a Celto-Germanic genetic subgroup (pace Mansion 1912) and, obviously, the existence of Italo-Celtic need not be refuted on this basis; cf. also that the list of NW IE innovations compiled by Oettinger (2003) contains not a single Celto-Germanicism. Instead, the situation presented here seems to reflect contacts between speakers of the IE dialects that later evolved into Proto-Celtic and Proto-Germanic. Religion and warfare seem to have been of particular concern.

Linking reconstructed prehistoric languages to archaeological findings is always risky business, but we may tentatively fix this cultural unity in time and space in Eastern Central Europe around 2000 BCE, when the pre-Celtic Ūnětice culture in the present-day Czech Republic bordered late, possibly pre-Germanic, varieties of the Corded Ware culture<sup>29</sup>. This scenario is at least partly compatible with conclusions reached by Kristiansen & Larsson 2005 and Kristiansen 2009: They envisage contacts between Pre-Germanic peoples and Pre-Celts immigrating from the South, spreading out over W Europe 2500-2000 B.C., not least by means of warfare and horses, until more hierarchical societies arise in the second millennium B.C.

## 5 A Fennic connection?

Most of the items in question look old and probably represent regional IE innovations, while others may have been taken over from the same third source. Interestingly, some of them seem to be shared with Balto-Fennic languages, suggesting a larger cultural continuum stretching further to the North. Particularly intriguing are Fi. *hepo*, *hevonen*, Est. *hobune* ‘horse’, Fi. *ratsu* ‘riding-horse’ and *kavio* ‘hoof’ (dial. *kapja*) since they all look Indo-European, but at the same time do not show the regular sound substitutions displayed by any attested Indo-European branch. Fi. *luppo* ‘lichen’ is inherited from Proto-Uralic, so if it is con-

<sup>29</sup> Ūnětice bodies are typically buried with jugs – meaning that *\*gan(dʰ)-*, too, could justifiably be categorized as belonging to the religious vocabulary.

nected to item no. 22, it must have been borrowed from Fennic into Celtic and Germanic.

If Balto-Fennic belongs to this cultural continuum, the question arises whether lexical exchange has taken place directly between Late Proto-Fennic and Pre-Proto-Celtic, or whether Pre-Proto-Germanic was always the provider: PCelt. *\*sanesto-* ‘secret advice’ (Matasović 2009: 322) is suspiciously reminiscent of Fi. *sanasto* ‘list of words’ (synchronically analyzable as *sana* ‘word’ + collective *-sto*), cf. the semantics of PCelt. *\*rūno-* and PGmc. *\*rūna-* (item 6 above) which in itself must be identical to Fi. *runo* ‘song; poem’. The vowel in *runo* is unexpectedly short, i.e. it does not behave as loanwords from Proto-Germanic normally do and may have been borrowed at an earlier stage. Mod. Ir. *lón*, pl. *lóinte* (> Eng. *lunch*) could represent Late Proto-Fennic *\*louna* ‘southwest; noon; lunch’ (Fi. *lounas*) which is derived from Proto-Uralic *\*luwe* ‘south’. Note that this word is already known to have been borrowed into Baltic (Latv. *launags* ‘lunch’, Lith. *láunagas* ‘dinner’). Fi. *maa* ‘land’ and its Balto-Fennic cognates go back to Proto-Uralic *\*maye*, reminiscent both in form and semantics of PCelt. *\*magos* ‘plain, open field’ > OIr *mag* ‘plain’, W *ma* ‘place’, Gaul. PN (*Arganto-*)*magus*), cf. Schrijver 2001:423. Fi. *tuoni* ‘dead’ < Late Proto-Fennic *\*tōne* could formally represent Proto-Celtic *\*doueno-* (item 27). Fi. *kalma* ‘grave; disease, Death-goddess, guardian of the abode of the dead’ could belong with PCelt. *\*klamo-* ‘grave’ (item 17)

For the same concept, PGmc. *\*haljō-* f. can be reconstructed (cf. e.g. ON *Hel* ‘death goddess’). It is most often seen as reflecting PIE *\*k<sub>1</sub>ol-<sub>1</sub>eh<sub>2</sub>*, derived from *\*kel-* ‘to cover, conceal’. However, if Fi. *Koljo* ‘name of a giant’ is a Germanic loan (IEW 553-554), the Finnish vocalism constitutes a problem—why is PGmc. *\*-a-* substituted with *-o-*? Moreover, a Proto-Finno-Ugric form *\*kolja* can be reconstructed also on the basis of Komi *kul* ‘water spirit’ and Mansi (Pelym dial.) *kuľ-nājər* ‘master of the netherworld, devil’. This word is internally analyzable as a participial form or agent noun derivative consisting of the Proto-Uralic verbal root *\*kole-* ‘to die’ and the agent-marker *-ja* with root-final *-e* regularly being dropped when a suffix is added<sup>30</sup>. Formally, nothing speaks against this word being a borrowing in the reverse direction, from Proto-Fennic into Pre-Proto-Germanic, i.e. at a stage before the Germanic sound shift and the development of *\*o* > *\*a*.

<sup>30</sup> For a slightly different analysis of the Uralic word, see Katz (2003:183).

## 6 Conclusions

An analysis of the NW Indo-European lexical material shared by Celtic and Germanic only is suggestive of the following scenario: The precursors of Celtic and Germanic evolved from different Indo-European dialect groups. Shortly after their migrations into Europe they came to form part of a cultural community, possibly influenced by indigenous populations or migrators from elsewhere. This had a significant impact on specific parts of the vocabulary, notably terms for religion and warfare. New derivatives were formed on the basis of Indo-European material, while some of the old ones were preserved in this area only. Some shared loanwords can be traced back to Late Proto-Fennic, spoken in their Northern vicinity. There are even indications that Late Proto-Fennic may have been in direct contact with Pre-Proto-Celtic, not always with Pre-Proto-Germanic as the provider. Celtic and Germanic peoples continued to influence each other, linguistically and in other respects, as they gradually developed the characteristics by which we define them.

## References

- De Bernardo Stempel, Patrizia, 2001: "Gotisch *INWEITþ guþ* und gallisch *ANDE-dion UĒDIFU-mi* (Chamalières, Z. 1)". – *Historische Sprachforschung* 114: 164-170.
- Birkhan, Helmut, 1967: "Das gallische Namens-element \**cassi-* und die germanisch-keltische Kontaktzone". – Wolfgang Meid (ed.): *Beiträge zur Indogermanistik und Keltologie*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck. Pp. 115-145.
- Birkhan, Helmut, 1970: *Germanen und kelten bis zum Ausgang der Römerzeit*. Vienna.
- Birkhan, Helmut, 1999: "Ein Strauß nicht durchwegs bekömmlicher Kräuter aus dem keltischen und germanischen Altertum". – Peter Anreiter & Erzsébet Jerem (eds.): *Studia Celtica et Indogermanica. Festschrift für Wolfgang Meid zum 70. Geburtstag*. Budapest: Archaeolingua. Pp. 43-52.
- Bjorvand, Harald, & Fredrik Otto Lindeman, 2000: *Våre Arveord*. Oslo: Novus.
- Campanile, Enrico, 1970: "Sulle isoglosse lessicali celto-germaniche". – *AION* 9: 13-39.
- Casaretto, Antje, 2004: *Nominale Wortbildung der gotischen Sprache*. Heidelberg: Winter.
- Chemodanov, Nikolaj Sergeevich, 1962: "Germano-kel'tskie leksičeskie paralleli". – *Sravnitel'naja grammatika germanskix jazykov* 1: 73-105. Moscow.

- Cohen, Paul S., & Adam Hyllested, 2007: "Monophthong for Expected v-diphthong in Greek". – Coulter George, Matthew McCullagh, Benedicte Nielsen, Antonia Ruppel & Olga Tribulato (eds.): *Greek and Latin from an Indo-European Perspective* [= Cambridge Classical Journal, supplementary volume 32], Cambridge: Cambridge University Press. Pp. 9-18.
- Elston, C.S., 1934: *The Earliest Relations between Celts and Germans*. London: Methuen.
- Hamp, Eric P., 2008: "Germanic \**fapmaz* and Gravity in the North". – *Early and Pre-historic Language Development in North-Western Europe* [= NOWELE 54/55]. Pp. 349-352.
- Hyllested, Adam, 2010: "The Precursors of Celtic and Germanic". – In Jamison, Melchert & Vine (eds.) 2010: 107-128.
- IEW = Julius Pokorny: *Indogermanisches Etymologisches Wörterbuch* I-II. Bern: Francke, 1959.
- Jamison, Stephanie W., H. Craig Melchert & Brent Vine (eds.): *Proceedings of the 21<sup>st</sup> Annual UCLA Indo-European Conference*. Bremen: Hempen.
- Jørgensen, Anders Richardt, 2006: "Etymologies to Go: Some Further Reflexes of Celtic \**keng*-.". – *Keltische Forschungen* 1: 59-72.
- Jørgensen, Anders Richardt, 2010: "Palatalization of \**sk* in British Celtic". – Benedicte Nielsen, Thomas Olander, Birgit Anette Olsen & Jens Elmegård Rasmussen (eds.): *The Sound of Indo-European*. Copenhagen: Museum Tusculanum Press. Pp. ??-??.
- Karsten, Torsten Evert, 1927: *Germanerna*. Stockholm: Natur och Kultur.
- Katz, Hartmut, 2003: *Studien zu den älteren indogermanischen Lehnwörtern in den uralischen Sprachen*. Heidelberg: Winter.
- Kerkhof, Peter Alexander, 2012: *Suffix variation in the PGmc. l-suffixes and the ablaut of the PIE l-stems*. MA dissertation from Leiden University.
- Kloekhorst, Alwin, 2008: *Etymological Dictionary of the Hittite Inherited Lexicon*. Leiden / Boston: Brill.
- Krahe, Hans, 1954: *Sprache und Vorzeit*. Heidelberg: Quelle and Meyer.
- Kristiansen, Kristian & Thomas Larsson, 2005: *The Rise of Bronze Age Society. Travels, Transmissions, and Transformations*. Cambridge: Cambridge University Press.
- Kristiansen, Kristian, 2009: "Proto-Indo-European Languages and Institutions: an Archaeological Approach". – Marc Vander Linden & Karlene Jones-Bley (eds.): *Departure from the Homeland: Indo-Europeans and Archaeology*. Washington DC: Institute for the Study of Man. Pp. 111-140.
- Kroonen, Guus, 2013 forthc.: *Etymological Dictionary of Proto-Germanic*. Leiden: Brill.
- Lane, George S, 1933: "The Germano-Celtic Vocabulary". – *Language* 9: 244-264.
- Lehmann, W., 1907: "Ahd. Widillo = ir. fiothal". – *Zeitschrift für Deutsche Wortforschung* 9: 313-314.
- LIV<sup>2</sup> = Helmut Rix (ed.): *Lexikon der Indogermanischen Verben*, 2nd ed., 2001. Wiesbaden: Reichert.
- Mallory, James P., & Douglas Q. Adams, 2006: *The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World*. Oxford: Oxford University Press.

- Mansion, Joseph, 1912: "Kelten en Germanen". – *Verslagen en Meededelingen der Koninkrijke Vlaamsche Academie* 1912: 1292-1308.
- Matasović, Ranko, 2009: *Etymological Dictionary of Proto-Celtic* [= Leiden Indo-European Etymological Dictionary Series 9]. Leiden: Brill.
- McCone, Kim, 1987: "Hund, Wolf und Krieger beim den Indogermanen". – Wolfgang Meid (ed.): *Studien zum indogermanischen Wortschatz*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck. Pp. 101-154.
- McCone, Kim, 1999: "OIr *erbaid* 'entrusts', *orb* 'heir' and *orbae* 'inheritance'". – Peter Anreiter & Erzsébet Jerem: *Studia Celtica et Indogermanica. Festschrift für Wolfgang Meid zum 70. Geburtstag*. Budapest: Archaeolingua. Pp. 239-242.
- McCone, Kim, 2005: "Mögliche nicht-indogermanische Elemente in den keltischen Sprachen". – Gerhard Meiser & Olav Hackstein (eds.): *Sprachkontakt und Sprachwandel. Akten der XI. Fachtagung der Indogermanischen Gesellschaft, 17.-23. September 2000, Halle an der Saale*. Wiesbaden: Reichert. Pp. 395-435.
- Mees, Bernard, 1998: "Celtic Influence in the Vocabulary of Hierarchy during the Common Germanic Period". – *Zeitschrift der Savigny-Stiftung der Rechtsgeschichte. Germanistische Abteilung* 115: 361-388.
- Meid, Wolfgang, 1965: "Zu einigen keltischen und germanischen Bildungen mit *st*-suffix". – *Die Sprache* 11: 122-130.
- Meid, Wolfgang, 1991: *Aspekte der germanischen und keltischen Religion im Zeugnis der Sprache*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Neu, Erich, 1995: "Hethitisch *tagu-* „dick, angeschwollen“". – *Historische Sprachforschung* 108: 1-5.
- Oettinger, Norbert, 2003: "Neuerungen in Lexikon und Wortbildung des Nordwest-Indogermanischen". – Alfred Bammesberger & Theo Vennemann: *Languages in Prehistoric Europe*. Heidelberg: Winter. Pp. 183-193.
- Olsen, Birgit Anette, 1988: *The Proto-Indo-European Noun Suffix \*-tlo- and Its Variants*. Copenhagen: Munksgaard.
- Polomé, Edgar C., 1983: "Germano-Celtic Isoglosses (Revisited)". – *Journal of Indo-European Studies* 11: 281-298.
- Polomé, Edgar C., 1994: "A Few Notes about Religious Correspondences between Celtic and Germanic". – Roland Bielmeier & Reinhard Stempel (eds.): *Indogermanica et Caucasica. Festschrift für Karl Horst Schmidt zum 65. Geburtstag*. Berlin: de Gruyter. Pp. 138-148.
- Puhvel, Jaan. 1991. *Hittite Etymological Dictionary*. Volume 3: Words beginning with H. Berlin / New York: Mouton de Gruyter.
- Rasmussen, Jens Elmegård, 1987 [1999]. "Miscellaneous Morphological Problems in Indo-European Languages IV, 22.: Greek *ἐρωτάω* 'ask, pray' – with a Note on the Indo-European Factitive Iteratives". – *APILKU* 6: 145-156 [Reprinted in Jens Elmegård Rasmussen, 1999: *Selected Papers in Indo-European Linguistics*, I: 309-311].
- Rasmussen, Jens Elmegård, 1989: *Studien zur Morphophonemik der indogermanischen Grundsprache*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.

- Rasmussen, Jens Elmegård, 1998: "Wege der Kulturverbreitung im Lichte der Wortformen". – Wolfgang Meid (ed.): *Sprache und Kultur der Indogermanen. Akten der X. Fachtagung der Idg. Gesellschaft*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck. Pp. 547-556.
- Ringe, Don, 2006: *From Proto-Indo-European to Proto-Germanic* [= A Linguistic History of English, vol. 1]. Oxford: Oxford University Press.
- Rübekeil, Ludwig, 2001: "Einige Bemerkungen zur Wortfamilie um germ. \*hatis-". – Béla Brogyányi (ed.): *Germanisches Altertum und christliches Mittelalter. Festschrift für Heinz Klingenberg zum 65. Geburtstag*. Hamburg: Kovač. Pp. 239-294.
- Rübekeil, Ludwig, 2002: *Diachrone Studien zur Kontaktzone zwischen Kelten und Germanen* [= ÖAW, Phil.-hist. Kl., Sitzungsberichte, 699]. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Schmidt, Karl Horst, 1984: "Keltisch und Germanisch". – Jürgen Untermann & Béla Brogyányi (eds.): *Das Germanische und die Rekonstruktion der indogermanischen Grundsprache: Akten des Freiburger Kolloquiums der Indogermanischen Gesellschaft, Freiburg 26-27 Februar 1981* [= Current Studies in Linguistic Theory 22]. Amsterdam: Benjamins. Pp. 113-153.
- Schmidt, Karl, Horst, 1986: "Keltisch-germanische Isoglossen und ihre sprachgeschichtlichen Implikationen". – Heinrich Beck (ed.): *Germanenprobleme in heutiger Sicht* [= Reallexikon der germanischen Altertumskunde, Ergänzungsband 1]. Berlin: de Gruyter. Pp. 231-247.
- Schmidt, Karl Horst, 1987: "Handwerk und Handwerker im Keltischen und Germanischen: Beiträge zu einem historischen Vergleich". – Wolfgang Meid (ed.): *Studien zum Indogermanischen Wortschatz*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck. Pp. 265-285.
- Schrijver, Peter, 2001: "Lost Languages in Northern Europe". – Christian Carpelan, Asko Parpola & Petteri Koskikallio (eds.): *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations* [= Mémoires de la Société Finno-Ougrienne 242]. Helsinki: Société Finno-Ougrienne. Pp. 417-425.
- Schrijver, Peter, 2004: "Apes, Dwarfs, Rivers, and Indo-European Internal Derivation". – Adam Hyllested, Anders Richardt Jørgensen, Jenny Helena Larsson & Thomas Olander (eds.): *Per Aspera ad Ateriscos. Studia Indogermanica in Honorem Jens Elmegård Rasmussen anno MMIV*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck. Pp. 507-511.
- Schumacher, Stefan. 2007: "Die deutschen und die Nachbarstämme: Lexikalische und strukturelle Sprachkontaktphänomene entlang der keltisch-germanischen Übergangszone". – *Keltische Forschungen* 2: 167-208.
- Stalmaszyk, Piotr, & Krzysztof Tomasz Witczak, 1991-1992: "The Celtic word for 'oats, Avena sativa', and its Indo-European equivalents". – *Lingua Posnaniensis* 34: 83-87.
- Stifter, David, 2009: "The Proto-Germanic Shift \*ā > \*ō and Early Germanic Linguistic Contacts". – *Historische Sprachforschung* 122: 268-283.
- Vine, Brent, 2002: "Att. ἐποτάω, Ion. εἰποτάω 'ask'". – *Glotta* 78: 203-221.
- de Vries, Jan, 1960: *Kelten und Germanen*. Bern: Francke.
- Watkins, Calvert, 1976: "The Etymology of Irish dúan". – *Celtica* 11: 270-277.

- Watkins, Calvert, 1991: "A Celtic-Latin-Hittite Etymology". – Tsvi Abusch, John Huehnergard & Piotr Steinkeller (eds.): *Lingering over Words: Studies in Ancient Near Eastern Literature in Honor of William Moran*. Pp. 451-453.
- Watkins, Calvert, 1995: *How to Kill a Dragon*. Oxford: Oxford University Press.
- Zair, Nicholas, 2012: Review of Jamison, Melchert & Vine (eds.) 2010. – *Kratylos* 57: 78-83.





## Again on Pigs in Ancient Europe: The Fennic Connection<sup>1</sup>

### Abstract

Proto-Celtic *\*mokku-* ‘pig’, *\*sukko-* ‘sow’ and *\*turko-* ‘wild boar’ are borrowed from Proto-Fennic where they are analyzable as inherited formations. Other Northern European terms can be shown to have an Indo-European origin: Welsh *cranan* ‘wild sow’, OIr *cráin* ‘sow’ belong to an earlier layer comprising both Germanic (OLFr. *chranni-chaltia* ‘pig’s den’) and Baltic (Lith. *šeřnas* ‘wild boar’), while Fi. *karjas* ‘wild boar’ is borrowed from an otherwise unattested PGmc. *\*garjaz* corresponding to Gk. χοῖρος, Alb. *derr*, from PIE *\*ǵhor-jo-s*. Latv. *cūka* ‘pig’ is not related to Lith. *kiaulė* ‘pig’ as usually assumed, but borrowed from PFc. *\*tsuka* ‘pig’ (> Fi. *sika*, Karel. *čugu* → N Saami *sokki* id.). NW PIE *\*porko-* ‘pig(let)’ is identified as an Altaic newcomer to the NW IE area on the basis of its widespread irregular variation in both IE and Fenno-Ugric, and the similarity with European words for ‘badger’, an animal typologically often compared to pigs. The lessons to be drawn are significant both culturally and linguistically: The great importance that boars played in Celtic and Germanic mythology must have been preceded by a centre of cultural gravity further to the North.

### 1 Hyonyms in Celtic – substratum material or Fennic loans?

It is well known that the wild boar played a significant role in ancient Celtic and Germanic (as well as in ancient Greek) mythology. Hamp (1987) has argued that the importance of boars and pigs went back to a

---

<sup>1</sup> The greater part of this article will be published in Birgit Anette Olsen & al. (eds.): *Etymology and the European Lexicon*, proceedings from *Fachtagung der Indogermanischen Gesellschaft* in Copenhagen (September 2012).

or North or Central European pre-Indo-European substratum<sup>2</sup>. A number of non-Indo-European pig terms, which we may conveniently name “hyonyms”, are of obscure origin, and many look non-Indo-European. For Celtic alone, Hamp lists the following as unexplained:

- 1 \**mokku*- ‘swine’ (OIr *mucc*, W *moch*, Gaul. PN *Moccus*, a swine divinity)
- 2 \**sukko*- ‘sow’ (W *hwch*, OIr *socc*-)
- 3 \**turko*- ‘(wild) boar’ (W *twrch*, OIr *torc*)
- 4 \**banuō*- ‘young pig’ (W *banw*, OIr *banb*)
- 5 W *cranan* ‘wild sow’ (White book of Rhydderch), OIr *cráin* f. ‘sow’
- 6 OIr *mat*, *mata* f. ‘pig’
- 7 OIr *cribais*, *cribu(i)s* ‘pig’
- 8 OIr *fithend* m. ‘boar (?)’

#### 1.1 PCELT. \**MOKKU*- ‘SWINE’

Hamp (1987: 187) began his overview with Proto-Celtic \**mokku*- ‘swine’, which is reconstructed on the basis of OIr *mucc* f. (originally a u-stem), W *moch* (collective; singulative *mochyn*), Breton *moc’h* (collective; singulative *penmoc’h*), Corn. *mōgh*, late *mōw* ‘id.’, and Gaul. *Moccus*, the name of a pig divinity. Hamp defined it as “perhaps the most prominent term ... notably lacking in IE cognates, a striking fact for the most pervasive generic lexeme for the pig”. He did not mention MLG and MDu. *mocke* f. ‘sow’, but as rightly stated by Kroonen (2013, pace Matasović 2009: 274-275) these Germanic forms are most likely to be loanwords from Gaulish and not directly from a third, unrelated source, since they

<sup>22</sup> Hamp (1987: 187) supposes that the cultural importance of the pig in Ancient Greece goes back to the same substratum, seeing that “Pre-hellenic was noot a satem language in type ... [and] is to be classed among IE dialects with the North European group ... We must look, therefore, for a pre-Greek movement of Indo-Europeans into the Aegean from the North, from as least as far North as Central Europe. It is clear then that the IE Prehellenic speakers could have brought with them to Greece the North European cultural values and institutions relating to the pig, these later to be incorporated into the Eleusinian mysteries”. He further ascribes the irregular variant *σῶς* (next to the regular *ῥς*) to this Prehellenic IE language. Fascinating as this scenario may sound, the prehistory and shaping of Greek culture goes beyond the scope of this article and will not be treated here.

are confined to the Southwestern (Franconian) part of the Germanic area. Hamp credited the substratum language even for the source of the u-stem formation, according to him typical of substratum words in Celtic, and found in one more pig-term, Welsh *hob* (see further under \**sukko*- below).

Remarkably in the light of Hamp's scenario, very similar terms specifically for 'sow' occur elsewhere in Northern Europe, namely among the Balto-Fennic languages. On the basis of Finnish *emakko*, Karelian *emakko*, *emäkkö*, Olonets *emäččü*, Lude *emäčču*, Estonian *emak*, Votic *emakko*, *emikko* we are able to reconstruct Balto-Fennic \**emakko*, \**emäkkö*<sup>3</sup> while Fi. *emokki* reflects BF \**emokke*. Following Hamp's line of reasoning, it would be natural to conclude that the North European substratum terms for pigs then stretched all the way to the Balto-Fennic area. Crucially, however, these words are internally analyzable within Balto-Fennic itself, even partly synchronically in the individual languages, as perfectly normal derivatives with the denominal suffix *-kko* from the noun *emä* 'womb; mother (also of animal)'. The Finnish variant with *-okki* (which, if regular, would go back to Balto-Fennic \**-okke*) has been formed to *emo* 'mother of animals, dam', in itself formed with a frequent denominal and deverbal suffix *-o* and common in Finnish compounds (e.g. *emo-lehmä* 'calver, mother cow', *emo-yhtiö* 'parent company').

Since derivatives with \**-kko* to the stem \**emä* must go back to the Balto-Fennic protolanguage, I will propose that Proto-Celtic \**mokku-* is simply borrowed from Early Balto-Fennic.

Prehistoric Celtic is normally attached to the Urnfield Culture of Central Europe (1300-750 BC), and its descendants, the Hallstatt Culture (800-500 BC, continuing into the historic La Tène Culture). Candidates for "Pre-Proto-Celtic" cultures are the Únětice culture (2300-1600 BC) and the Tumulus culture (1600-1200 BC). Kristiansen and Larsson (2005) and Kristiansen (2009) place the emergence of Pre-Celtic culture in Western and Central Europe 2500-2000 BC. As I

<sup>3</sup> BF \**emakko* has undergone a common, but not entirely regular, morphophonemic development of *ä* > *a* / *eC\_C(C)o* (cf. e.g. Fi. *kesä* 'summer' ~ *kesakko* 'freckle', *elä-* 'to live' ~ *elanto* 'livelihood') while the variant \**emäkkö* displays the expected vowel harmony. *-ikko* (and *-ikkö*) as in Votic *emikko* mostly occurs after stems in *-e-* or *-jä-*, but sometimes even after *-ä-*stems, cf. Fi. *silmä* 'eye' ~ *silmikko* 'bud', *heinä* 'grass' ~ *heinikko* 'meadow'. These processes have been described by Campbell 1980 (257-258 with references) from a purely Finnish perspective, but they must be regarded as common Balto-Fennic phenomena because reflexes of all types occur throughout Balto-Fennic.

stated in Hyllested (2010) there are other lexical indications that Middle and Late Proto-Fennic<sup>4</sup>, spoken in the northern vicinity of these cultures, may have been in direct contact with Pre-Proto-Celtic, not always with Proto-Germanic as the provider<sup>5</sup>. Late Proto-Fennic (Proto-Balto-Fennic) is often assumed to have emerged around 1500 BC (see e.g. Janhunen 2009), preceded by the Middle and Early Proto-Fennic (Proto-Fenno-Saami) stages.

The most important argument for regarding *\*mokku-* as a loan from Fennic<sup>6</sup>, however, is the big lexicological picture: The fact that other Proto-Celtic hyponyms also have good candidates for a source in Balto-Fennic.

#### 1.2 PCELT. *\*SUKKO-* ‘SOW’ ?

Another well-known puzzle concerns PCelt. *\*sukko-* (W *hwch*, OBret. *hoch*, Mod.Bret. *houc’h*, Corn. *hoch*, OIr *socc-* ‘sow’) which looks intriguingly similar to the most common Indo-European hyonym *\*suH-s* (Lat. *sūs*, Gk. *ŷς*, Alb. *thi*, Toch. B *suwe*), but whose final part *-kk-* has been difficult to account for. The classical handbooks reconstructs a root variant *\*seu-k-* next to *\*seu-H-*, more or less explicitly interpreted

<sup>4</sup> I use Kallio’s (2007, 2012) trichotomy of Early, Middle, and Late Proto-Fennic. Of these, Early Proto-Fennic, the stage before any distinctively Balto-Fennic innovations, is the traditional name for Proto-Fenno-Saami (I hesitate to accept Kallio’s inclusion of Mordvin here), and Late Proto-Fennic equals Proto-Balto-Fennic, the protolanguage of all the Balto-Fennic languages (perhaps except South Estonian, see Kallio 2007). Important for our discussion, Middle Proto-Fennic is “the stage largely recoverable by internal reconstruction immediately before the development *\*ti > \*ci*” (Kallio 2012: 166, fn. 9).

<sup>5</sup> Among my proposals for direct loans from Proto-Fennic into Proto-Celtic are PFc.*\*sanasto* ‘list of words’ (from *\*sana* ‘word’ + loc.suff. *\*-sto*) → PCelt.*\*sanesto-* ‘secret advice’ (compare the semantics of the reverse loan Fi. *runo* ‘(traditional) song, poem’ < PGmc. *\*rūna-* or PCelt. *\*rūnō-* ‘secret’); *\*louna* ‘southwest; noon; lunch’ (from PU *\*luwe* ‘south’ + suff. *-na*) → NIr. *lón*, pl. *lóinte* ‘lunch’ (further borrowed into Eng. *lunch*, *luncheon*; also a Fc. loan in Baltic, Latv. *launags*, Lith. *láunagas* ‘dinner’); PFc. *\*maya* ‘land’ → PCelt. *\*magos* ‘plain, open field’; PFc. *\*kalma* ‘disease; grave; Death-goddess’ PCelt. *\*klamo-* ‘diseased, leprous’. PFc. *\*tōne* ‘dead’ may represent a reverse borrowing from PCelt. *\*doueno-* ‘(mortal >) man’ id. into PFc. (cf. PGmc. *\*dewena* ‘mortal’, *\*dawjan-* ‘to die’), parallel to the hyonyms. Cf. Hyllested 2010: 123–124.

<sup>6</sup> A hypothetical hybrid form *\*emokko* would appear even closer to the Celtic forms. The deletion of *-e-* seems to presuppose stress shift to the second syllable *-mók-*. We will deal with this below.

either as an old morphophonemic alternation or as an extension proper whereby an old “determinative” *-k-* has been added to the root in its strictest sense, i.e. *\*seuH*<sup>7</sup>. The extended version would then make up the protoform not only of Celtic *\*sukko-*, but also of Germanic *-k-* variants like OE *sugu*, OS *suga* ‘sow’, as well as Latin *sucula* ‘young sow’ and Skt. *sūkára-* ‘(male) boar’.<sup>8</sup>

Reconstruction of root variants and root extensions is problematic in general because it gives etymologists an extra chance to replace or add a root-final consonant whenever these final elements turn out not to match. The field of Indo-European studies have now reached a stage where scholars should try to either identify the source or function of such irregular and unexpected elements. If they are just conveniently reconstructed whenever a solution is needed, without an explanation of their origin or function, it jeopardizes our chances of staying on the right track and discover the actual conditions of the past. Unexplained root-final elements simply leave too many options open. Hamp (188) justifiably states “We have no license to drop the laryngeal in IE *\*suH-*, and in fact the claimed Indo-Iranian comparanda conserve the *\*ū*. I therefore see here a substratum *\*suk-*, geminated in Celtic”.

Even those scholars who accepted *\*seuk-* as a variant to account for PCelt. *\*sukko-* still have had difficulties explaining the Celtic gemination. Needless to say, “expressive gemination” has been among the suggestions (Polomé 1953: 541). Testen (1999) analysed the pig-names as original compounds where *-kku-* reflects PIE *\*p̥ku-* ‘livestock’ (the zero-grade of *\*péku* n.), but an animal name like *\*brokko-* ‘badger’ can hardly count as a term from the field of animal husbandry; besides, other stems than *u-* stems are found in the Celtic material.

Kroonen (2011) makes a successful case in trying to eliminate the PIE variant *\*seu-k-* altogether: He shows, first of all, that West Germanic forms with *-g-* (OE *sugu*, Mod.Du. *zeug*) simply owe this velar to a regular development of hiatus or *-w-* between two high vowels if at least one of them is *u*. Norwegian *sugge*, Sw.dial. *sågg* ‘sow’ would have arisen by normal verschärfung of *\*-ww-* to *\*-ggv-* in Nordic. Second, he points out that Lat. *sucula* is simply formed with the normal diminutive ending -

<sup>7</sup> Testen (1999: 191) states that “*\*sukko-* shows phonological problems that complicate any interpretations based upon its obvious similarity to Indo-European *\*sū-*”.

<sup>8</sup> The latter interpretation is problematic seeing that no trace of length is left in Latin, Celtic nor Germanic; all forms in these languages begin with *\*sū-* and cannot reflect *\*suH-*.

*cula* that is added to vocalic stems (cf. *auri-cula* ‘little ear’, *avi-cula* ‘little bird’, *api-cula* ‘little bee’); a *\*\*sū-(u)la* would be impossible, and the only way to form a diminutive with *-la* to *sūs* would be to add *-cula* to the stem. Hence, there is no need to believe that *-c-* belongs to the root in the Latin case either. Third, as already suggested by Fick, Falk & Torp (1909), Skt. *sū-kāra-* m. ‘wild boar’ is probably simply a compound, meaning literally ‘pig-reproducer’<sup>9</sup>. Kroonen’s conclusion is that PIE *\*seuk-* did not exist, since only Celtic *\*sukko-* cannot be explained away and needs a source<sup>10</sup>. His solution is to assert for the Celtic animal names in *-kko* a Germanic source where *-kk-* derives from n-stems to roots in *-k-* via Kluge’s Law. However, of the Celtic animal names, only PCelt. *\*bukko-* ‘billy-goat’ has a safe counterpart in Germanic.

Besides, there is one more Indo-European term to take into account. Interestingly, PCelt. *\*sukko-* is somewhat reminiscent of Latv. *cūka* ‘pig’ whose etymology is also disputed: In native Latvian words *c-* usually occurs before *-e-*, *-i-* because it has developed from late palatalization, or it is a borrowing from Estonian, cf. Latv. *cīrele* ‘lilac’ < Est. *tsirel*, dial. for standard *sirel*. Jānis Endzelīns simply described *c-* in *cūka* as irregular from *\*sūka* which he equates with the Germanic, Latin and Indic forms in the previous paragraph (Kaspars Ozoliņš, p.c.). Karulis (1992) prefers to group *cūka* with Lith. *kiaūlė* ‘pig’, visualizing a zero-grade of *\*keu-/kū-*, cf. *kaūkt* ‘yell; howl’, and various toponyms such as *Kūkas*, presumably ‘place with a lot of wild pigs’. This would imply an analogically mixed root where the original PBalt. distribution *\*kiau-/kū-* was analogically levelled to *\*kiau-/kiū-*. While not impossible, the Latvian and Lithuanian words actually do not have that much in common, and as shown by Hamp (1986), the original meaning ‘pig’ was probably connected specifically to a stem containing the *-l-*. Hamp equates *kiaūlė* with the element *Cul-* in the Welsh PN *Culhwch*, referring to a divine pig, a cousin of Arthur, of the same class as *Twrch Twyth*. This means

<sup>9</sup> Formally corresponding to the Middle Persian proper name *Hukar* < Proto-Iranian *\*hūkara-* (Blažek 2010: 90)

<sup>10</sup> Middle Persian *xūk*, Modern Persian *xūg* (Blažek 2010: 88, 90), not mentioned by Kroonen, probably derives from a typical secondary formation in Iranian, *\*hū-ka-*, whereby the suffix *\*(i/a)ka* is added to stems that would otherwise be very short. Alternatively, it may have been a diminutive formation denoting the piglet, cf. e.g. Alb. *derk* ‘piglet’ < Proto-Alb. *\*dar-ika* next to *derr* ‘pig’ (Orel 1998: 61). Other Iranian forms show expected reflexes of the root-noun, e.g. Young Avestan *hū-* ‘pig’, Ossetic (Digor) *xu*, (Iron) *xʷy* id. Laconian Greek *σίκα* (? < *\*tika*), likewise absent in Kroonen’s account, may simply be onomatopoeic (Katz 2003: 206–207), cf. also Polish *dzik* ‘pig’.

that Culhwch is basically a tautological formation, simply meaning '(the) pig'. Hamp reconstructs *\*keuliā*, not as a PIE form, but rather as a form borrowed from a North European substratum language into at least Celtic and Baltic. He includes in this word-family Lith. *kuilys* 'boar', strikingly reminiscent of Culhwch's father's name, *Cilydd* < *\*kūlios*.

Another possibility for Latv. *cūka* would be to assert an onomatopoeic origin parallel to that behind Pol. *dzik* 'pig', a sound sequence which seems to be geographically widespread as a word used for attracting the pigs. However, such words often seem to originate from real nouns, cf. e.g. Ukr. *gus*, a word used for calling geese, and Da. *hyp*, a word used for making horses move (cf. *hoppe* 'mare' < *\*huppan-*).

In this particular case, there is no need to go far for a source. I believe that the obvious source of Latv. *cūka* is the term for the same animal in the languages spoken to the north of Latvian, namely the Balto-Fennic languages. The situation within Balto-Fennic itself is unclear, too, seeing that while Finnish *sika* reflects *\*sika*, some of its dialects have *tsika*, pointing rather to Balto-Fennic *\*tsika*; Karelian *čugu* meanwhile points to *\*tsuka* (BF *\*suka* would have yielded Karelian *\*šugu*), and borrowings into Saami such as N Saami *sokki* point to a fourth variant, *\*suka*. This messy situation has led some Fennicists to believe that we are dealing with different etyma, but this is too hasty a conclusion. First of all, the irregular vocalism, although its distribution is unaccounted for, is not unparalleled: for example, it makes little sense to separate BF *lintu* 'bird' from FU *\*lunta* 'bird' (cf. N Saami *loddī*), and this word also shows up with *-i-* for expected *-u-* in Fennic only. The situation is admittedly not exactly parallel since, unlike the pig-word, the bird-word does not have alternating vocalism *within* Balto-Fennic, but in fact this exactly speaks for uniting the pig-words since in their case we are not only conjecturing the occurrence of a *-u-*; we see it attested in front of our eyes. In other words, if we accept *lintu* < *\*lunta* in the first place, there is no reason to doubt *\*(t)sika* ~ *\*(t)suka* < *\*tuka* on the basis of the vocalic variation. The consonantal variation in the onset is pretty straightforward: *ts-* is not phonotactically allowed in Finnish, and although the development of Fennic-Saami (Early Proto-Fennic) *\*ti-* > *\*si-* and *\*tū-* > *\*sū-* is not regularly extended to the third high vowel, one might still visualize a few cases where this tendency initiated, especially if there was a dialectal variation between *\*tsi-* (maybe still *\*tsū-* at the time) and *\*tu-*. Besides, the Fenno-Volgaic reconstruction is actually *\*tuka* as can be seen from Mordvin *tuvo*, so the Balto-Fennic onset al-



ready presents an irregularity that needs to be explained, regardless of the situation within Balto-Fennic.

These observations mutually confirm each other: While the reconstruction of a BF *\*tsuka* renders it possible to provide a straightforward source for Latvian *cūka* ‘pig’, correspondingly the identification of Latvian *cūka* as a Balto-Fennic loan confirms the BF reconstruction with *\*ts-*. Incidentally, both elements of the Latv. compound *mežacūka* ‘wild boar’ therefore have the same etymological source as those in Fi. *metsäsika* ‘badger’ (since *metsä* ‘forest’ is a borrowing from Baltic *\*medīa-* ‘id.’).

Despite the obvious chronological differences between the situation for Latvian *cūka* and PCelt. *\*sukko-*, I believe that the most probable origin for the latter is also Balto-Fennic *\*tsuka*. Note that while Latvian shows a secondary lengthening in its reflex, Saami *sokki* has a geminate consonant. These different kinds of lengthenings probably reflect different attempts to render what was heard as a kind of heavy syllable in the target language. As a parallel, Modern Fennic words of the structure CVCV (e.g. *sika*) are frequently perceived by e.g. speakers of Danish as having an unexpectedly long (or partly stressed) second syllable, since the vowel of this second syllable is longer than in Danish words of the same structure (e.g. *mokka* ‘[café] mocha’, *sikahjort* ‘sika deer’), making it sound like a compound to Danes. My guess is then that both the Balto-Fennic consonantal alternation in paradigms and the persistent stress on the first syllable had some effect on the foreign renderings of Fennic *\*tsuka*, resulting in a geminate in Saami *sokki* and PCelt. *\*sukko-* but as vowel length in (the probably younger, but not necessarily very young) loan in Latvian, *\*cūka*. The Balto-Fennic *-k-* was thus only rendered by a geminate consonant in the languages where such consonants existed, whereas Latvian, not possessing geminates, expressed the length in the preceding vowel instead. One may also note that secondary lengthening of both vowels and consonants is pretty commonplace in Fennic itself, even in inherited words.

The u-stem formation in *\*mokk-u-* and *\*sukk-u-o-* (> W *hob*) that Hamp (1987) identifies as part of the substratum features do not seem to be detectable in Fennic; the Karelian *-u* is regular from *-a*, and *\*tsuka* looks like a normal Fennic a-stem.

## 1.3 PCELT. \*TURKO- ‘WILD BOAR’

A third puzzling term is PCelt. \*turko- (W [obs.] *twrch* ‘pig’ m., pl. *tyrch*(-ot), OCorn. *torch*, OBret. *torch*, MBret. [Catholicon] *tourch* ikd., OIr *torc* [masc. o-stem] ‘wild boar’). It is remarkable that even this term ends in -ko. The sequence -Rkk- was non-existent in Proto-Celtic (the Brythonic development to -ch is regular after resonant) so we can even define the last part of all three words as identical<sup>11</sup>. McCone (1992, 1993) identified the Celtic term with Avestan *θβərəsō* (occurring once in an Avestan fragment of the *Pahlavi Rivāyat accompanying the Dādestān-ī Dēnīg*), reconstructing PIE \*tʰorkós with an original meaning ‘cutter’ (referring to the boar’s “notorious talent for tearing and uprooting with his sharp tusks”) which has since been generally accepted (with some reservations Lubotsky 1994; Mallory & Adams 2006: 139). The word would then rhyme with \*porkós ‘(young?) pig’ and even with *ior̥kós* ‘deer’, seemingly revealing a structure for the formation of names for mammals (cf. \*h₂rt̥kós ‘bear; beast’, PGmc. \*elha- ‘elk’, \*selha- ‘seal’, e.g. Kroonen 2013 under entry \*baruga-).<sup>12</sup>

As already noted by Hamp (1989: 188), McCone’s reconstruction of an o-grade in \*tʰorkós is misleading since the attested Celtic forms unanimously point to a persistent zero-grade \*turkos. This is significant since, as later shown by Lubotsky (1994), no other ablaut grade is to be found in derivatives underlying this postulated root ‘to cut’, neither in Indo-Iranian nor in Greek which would be the two other branches to have allegedly preserved reflexes of it (however, in Greek and Indo-Iranian the vocalization is different, pointing to \*tʰrk- and not \*turk-)<sup>13</sup>. This similar behavior obviously speaks in favor of a connection between the Celtic boar-word and the ‘cut’-root.

Meanwhile, however, several factors speak against this. Not only is the attestation of Av. *θβərəsō* restricted to a hapax in a Pahlavi fragment;

<sup>11</sup> Matasović (2009) reconstructs PCelt. \*turkko-, but the development of PCelt. \*-k- to Welsh *ch* is regular in the position after -r-.

<sup>12</sup> Some of these examples are equally disputed; Schindler (1966) regarded \*selhaz as an Early Proto-Fennic loanword in Germanic, cf. LPFc. \*hūlyeh (Fi. *hylke*, gen. *hylkeen*) ‘seal’ < EPFc., PFU \*šülkeš or \*čülkeš.

<sup>13</sup> Av. *θβərəs-* ‘to cut, shape’, *upa-θβərəsqn* acc.pl. ‘hole, split’, *θβərəsāscā* acc.pl.m. ‘(an) end, split’, *θβōrāštar* ‘creator’ representing \*θβərəštar with a common coloring of ə to ō by a preceding (or following) labial; Ved. *Tváṣtar-* ‘the god-creator’, where the original *r̥* > *a* has become identical to the following vowel under dissimilatory influence of a second -r- later in the word; and Gk. σάρξ, Aeol., Dor. σῶρξ ‘flesh’.

its very existence is uncertain and it may in fact be an invention by the scribe as suggested by Hoffmann (1969: 35). As he writes, the passage *yaθa vā azō sacainiš yaθa huš θβərəsō*, functioning as the subject of a following phrase in Pahlavi meaning '(are) to be killed in the sacrifice for the gods', corresponds to an almost identical fragment in the Nīrangistān that reads *ḡaaθa vā azō scaēnīš ḡaθa huš pərəsō*; since the context is about sacrificial animals about to be slaughtered, it seems very likely that in the Rivāyat the obsolete *pərəsō* was reshaped under the influence of the verb *θβərəsaiti* 'to cut' (although of course the replacement would be natural if the two words were really synonyms in the first place). If *θβərəsō* is really a scribal innovation, the semantic bridge between Indo-Iranian and Greek 'to cut' on one side and Celtic 'boar' on the other is obviously destabilized.<sup>14</sup>

Furthermore, the rhyming word *\*porkōs* (> Av. *pərəsō*) as we shall see later, probably does not go all the way back to PIE and may be of non-Indo-European origin. Besides, whatever the reason for the all-dominant zero-grade in Greek and Indo-Iranian reflexes of *\*turk-*, an isolated view on Proto-Celtic *\*turkos* as a pig-name would result in an interpretation of the odd zero-grade as signs of a non-Indo-European origin (cf. Hamp 1989: 188; "the scarcely IE sequence *\*ur*").

Balto-Fennic again provides us with an unheeded candidate for a source which is remarkably close both semantically and formally: Karelian *torakko*, *torikko* means nothing less than 'tusk of wild boar', and this word, too, is an unproblematic internal formation within Balto-Fennic, namely as derivative of *tora*, also 'tusk of wild boar', with the denominal suffix *-kko*, cf. Fi. *tora-hammas* 'tusk' (*hammas* 'tooth'). The noun is widespread within Fenno-Ugric and identical to the verbal stem FU *\*torV-* (> e.g. Fi. *tora-*) 'to struggle, to fight, to battle' whose Saami reflexes such as N Saami *doarro-* and Lule Saami *tårrō-* specifically mean 'to fight with the horns; thrust (of mammals)', denoting actions typically carried out by boar tusks. Thus, while the ultimate underlying semantics may not be very far from what McCone suggested, the word's character as Proto-Fennic loanword seems clear to me especially in the light of similar scenarios for *\*sukko-* and *\*mokku-*.

<sup>14</sup> McCone was of course aware also of this part of Hoffmann's article, but writes (1992: 99): "Convincing though the interpretation of *pərəsō* as *parāsō* < *\*porkōs* is, its corruption to a non-existent *θβərəsō* somewhat reminiscent of the verb 'to cut' is a less attractive postulate".

## 2 Indo-European terms

### 2.1 PIE \**ĥ*<sub>ER-N</sub>- ‘WILD BOAR’ ← ‘TUSK; BRISTLE’?

PFP terms seem to confirm the cultural importance of the pig in prehistoric Northern Europe, even in cases where nothing points to a Fennic origin. Hamp (1987: 189) declares himself incapable of etymologizing OIr. *cráin f.* ‘sow’, but suggests Welsh *cran* ‘wild sow’ (attested in the White Book of Rhydderch and then once later) as cognate. An exact Proto-Celtic reconstruction is admittedly not possible on the basis of these two forms. However, the common Lithuanian word for the ‘wild boar’, *šeřnas*, with a variant *šernùkas*, is suggestive of a formation from the PIE stem \**ĥer-n-* ‘horn’, referring to the boar’s tusk; cf. the double meaning of Skt. *śṛṇ-gá-* ‘horn; elephant’s tusk’<sup>15</sup>. It is tempting to include the first member of OLFr. *chranni-chaltia* ‘pig’s den’ (Quak 1983)<sup>16</sup> whose first member has so far been considered obscure. A term \**ĥer-n-* designating the wild boar, thus seems to unite Celtic, Germanic and Baltic.

### 2.3 PIE \**G<sup>h</sup>ORĭO-* ‘WILD BOAR’

The form \**ĥer-n-* mentioned immediately above may appear similar to Fi. *karjas*, *karju* ‘wild boar’, *karja* ‘livestock’ (cf. also Est. *karjane* ‘shepherd’) but these Balto-Fennic forms notably lack the nasal element. They are conventionally connected to the Finnish verb *karjua* ‘to roar’ and the noun *karjainen* ‘rut, rutting (of male animals)’, but the question is in what direction the derivational process originally went. The ending *-as* in *karjas* suggests that we are dealing with an Indo-European loanword. A PIE \**ĝhór-ĭo-s* can indeed be reconstructed on the basis of Gk. *χοῖρος* ‘wild boar’ and Alb. *derr* (Mallory & Adams 2006: 142). The Proto-Albanian form was \**darja* (Orel 1998: 61), corresponding perfectly to the Greek form; an alternative shape \**ĝhóĭ-ro-s* (preferred by Demiraj 1997: 131–132) would admittedly also be possible on the basis of Greek and Albanian only if the original Alb. sg. \*\**darr* was generalized the umlauted

<sup>15</sup> Lith. *šeřnas* with *métatonie rude* also exists. Smoczyński (2003: 10, 92) prefers an inner-Baltic derivative from *šeriai* (also *šerys*) ‘bristle’. Cf. also Hyllested & Gliwa 2009: 50) on the mechanisms behind this derivational process.

<sup>16</sup> I thank Guus Kroonen for having drawn my attention to this form.

pl. *derra*<sup>17</sup>. However, note that the Fennic forms now confirm a PIE *\*ǵhór-io-s* if they are borrowings from an otherwise unattested, but perfectly matching, PGmc. *\*garjaz*. One might visualize that another kentum language than Germanic, known or unknown, could be the source of *karjas*, but it would have to be a language that lost the velar-palatal distinction since PIE *\*ǵh-* would otherwise be attested as a sibilant, cf. Fi. *salko* 'pole, stake' ← PIE *\*ǵh<sup>h</sup>alg<sup>h</sup>-*. In forms old enough to have retained an o-vocalism, palatals are usually substituted with sibilants in FU, cf. Fi. *koipi* 'leg of a bird; (colloq.) human leg' < (NW) PIE *\*kōipo-* 'pole, stake' (Skt. *śēpa-* 'tail, penis', Alb. *thep* 'peak, point, cog, tooth', Lat. *cippus* 'pole, stake')<sup>18</sup>, but the a-vocalism in *karjas* points to Germanic as the most probable source. This would, conveniently for our reconstruction of the prehistoric situation, mean that Indo-European borrowings of hyonyms into Fenno-Ugric took place much later than the borrowings of hyonyms in the reverse direction.

### 3 Motivation for borrowing

That Celtic and Germanic hyonyms that can be shown to have originated in Fennic should strike one as unexpected, partly since boars play a significant role in Celtic and Germanic (as well as in Greek) mythology, partly since wild boars are generally 'Southern' animals in Europe. However, there are chronological layers to distinguish: These terms must go back to a time from before the emergence of these specific traits in at least the "Celtic" (and probably also the "Germanic") cultures and religions. The terms themselves and the cultural significance they reveal must both emanate from a common non-Indo-European source. The question then remains if we can trace any extralinguistic evidence for a special importance of pigs among the Fennic peoples. Tacitus wrote on the Aestii, a Northeast European tribe in the Baltics:

"They worship the mother of the gods: as an emblem of that superstition they wear the figures of wild boars: this boar takes the place of arms or of any human protection, and guarantees to the votary of the goddess a mind at rest even in the midst of foes"

<sup>17</sup> Huld's (1984: 148) reconstruction of an unparalleled derivative *\*suojn-ro-* from the stem occurring in PGmc. *\*swina-* (PIE *\*suH-ih<sub>3</sub>no-*) is thus unnecessary.

<sup>18</sup> In fact, the Fennic form confirms IE o-vocalism which is not otherwise directly attested in this form.

While the name of the Aestii (also *Aestiorum gentes*) is no doubt the source of that of Estonia, it has generally been assumed that the Aestii were in fact speakers of Baltic and not Balto-Fennic languages, whose name was later transferred to the Balto-Fennic Estonians. However, Bammesberger & Karaliunas (1998) convincingly show that the Aestii, and that the original ethnonym, definitely denoting a Fennic people from the point of view of the Scandinavians, had an extra *-r-* in stem (*Eistr-*)<sup>19</sup>. As I demonstrate elsewhere (Hyllested forthcoming), the Baltic stem *\*aistra-* had a meaning synonymous to that of PGmc. *\*finōn*-<sup>20</sup> and is in all probability a loan translation, clearly indicating that the name denoted Fennic peoples.

Archaeological evidence can be supplied. Sites from the Pitted Ware culture (3200-2200 BC) on southern Scandinavian coasts from Svealand and Åland to the Danish island of Funen contain pig bones in large quantities emanating from domesticated pigs rather than wild boar. It is known that they lived side by side with battle-axe peoples, traditionally attached to Indo-European and Pre-Proto-Germanic expansions. The people of the pitted ware were not direct ancestors of Northern Scandinavians, but more closely related to peoples of the contemporary Baltic region (Rowley-Conwy & Storå 2007; Malmström & al. 2009), and blending of styles and techniques between pitted ware and battle-axe peoples took place especially in the later half of the period, 2700-2200 BC (Larsson 2003). It is clear from the datings listed above that this period does not fit exactly with the Balto-Fennic protolanguage but rather with the traditional dating of Fenno-Volgaic or Fenno-Saami. But at least we have reason to believe that the cultural significance of the pig continued into the Balto-Fennic period, and perhaps, in the light of Tacitus' account, even into historical times.

<sup>19</sup> Cf. *Eistra dolgi* 'the Estonian enemy' (Ynglingatal) and *devenit in Eistriam, puer Olavis Eistriis in servum venundatur* (Historia Norvegiae). The original vocalism *\*aist-* is secured by Old Gutnish *utan foru i aina oy viþr Aistland, sum haitir Dagaþi* '[they] travelled to an island off Estonia called Dagö' (Guta Saga).

<sup>20</sup> 'Pimple' or other protuberances from the skin of humans or animals such as 'fish scale', 'fin', 'larvae under the skin of cattle (causing folly)'. The Baltic meaning is reconstructed on the basis of a loanword in Livonian *āistar* 'pimple' (the dialectal variant *vistar* can only be explained on the basis of Baltic) and Lith. *aistrà* 'intense passion'; cf. Gk. οἷστρος 'intense passion; larvae under the skin of cattle', ON *eistra* 'testicle', OHG *eiz* 'larvae under the skin of cattle', Gk. οἶδος 'tumor'.

Also seal bones are prominent among the findings in Pitted Ware sites; note that the dating fits perfectly with Schindler's etymology for PGmc. *\*selhaz* 'seal' mentioned in fn. 11 above.

#### 4 Badgers and pigs

It is relevant now to introduce PCelt. *\*brokko-* 'badger' (Ir. *broc*, W *broch*)<sup>21</sup>, into the discussion for two reasons: First, it is an animal name of obscure origin containing the geminate *-kk-* that we recognize from the pig terms. Second, it is typologically common to compare badgers to pigs and their offspring to piglets, cf. e.g. the aforementioned Fi. *metsäsika* 'badger' (lit. 'forest pig' ~ Latv. *mežacūka* 'wild boar'), Nw. *svin-toks* 'badger' (1<sup>st</sup> part *svin* 'swine', 2<sup>nd</sup> part *\*pahsu-* 'badger'), Eng. *sow*, Da. *so* 'female of pig', but also 'female of badger', Eng. *boar* 'male of pig', but also 'male of badger', and Da. *gris* 'piglet', but also 'young of badger, cub'.<sup>22</sup>

Hence we should consider the possibility that PCelt. *\*brokko-* could be derived from word for 'pig' of non-Celtic origin. It is almost for certain connected to W Germanic *\*brakka-* '(scent) hound, dog used for hunting' (OHG *bracko*, MDu. *bracke* 'sleuthhound') since hounds have been used for hunting game such as (very often specifically) badgers or boars since ancient times<sup>23</sup>, cf. also the modern term *dachshund* < NHG

<sup>21</sup> Borrowed into OE as *brocc*; from Eng. it has further been borrowed into Danish as *brok*, mostly used in the definite form *brokken* (*Brokken*) as a kind of semi-personal name, especially common in hunter's language.

<sup>22</sup> More examples can be found in Ritter (1975). We might add as a modern example the name of the 'hog badger' (*Arctonyx collaris*) from Southeast Asia. In this light, it might be worth investigating whether the otherwise opaque Eng. word *badger* could be a borrowing from Brythonic (cf. W *baedd*) and perhaps even confirm the reconstruction *\*badjo-* rather than the alternative *\*basjo-*. The root in Lat. *fodiō*, Hitt. *padda-* 'to dig (the ground), bury', Lith. *bedū* 'dig' comes to mind when you compare e.g. Danish *grævling* 'badger' ~ *grave* 'to dig' (and *grav* 'fox or badger earth, burrow'), The root is attested in Celtic but with innovative meanings (cf. Hyllested 2010: 115).

<sup>23</sup> I find Kroonen's (2013) alternative reason for linking the two words, that the badger is an "animal with a strong sense of smell" less relevant. His semantic reconstruction 'sleuth dog' is also somewhat anachronistic since he seems to prefer that the word is Proto-Germanic; 'sleuthhound' is admittedly the meaning in both Middle Dutch and Old High German, as well as in modern Dutch *brak* and modern German *Bracke* (hence Eng. *bracke*) but the breed does not go back to ancient times. An less specified meaning 'hound, hunting dog' or 'scent dog'

*Dachshund*, lit. 'badger dog'.<sup>24</sup> The origin of the latter breed is uncertain, but probably not bred as early as when Proto-West Germanic was spoken. Cf. also ON *Brokkr*, the name of the dwarf who creates the mythological boar *Gullinbursti*.

Ru. *barsuk* 'badger' is a loanword from Turkic *\*borsuk* (*\*borsuq*) id., derived internally in Turkic from *\*bor*, *\*boz* 'grey' and akin to Written Mongolian *borki* 'old badger' (Khalkha Mongolian *borx*, Buryat *burxi*, Kalmyck *bork* 'badger') borrowed into Tuvin (a Turkic language spoken to the North of Mongolian in Central Asia) as *murgu* 'male marmot' (Khabtagaeva 2009: 159). The word has also been borrowed into Hungarian as *borz* 'badger'. Fi. *mäyrä* comes from older *\*mäkrä*, cf. Karelian *mäkrä*, Est. *mäger*. The Balto-Fennic word can only be connected via dissimilation from *\*märkä*. Celtic *\*br-* can come from older *\*mr-*, and a stress shift with vowel loss in the first syllable is reminiscent of the hypothetical process *\*emokk-* > *\*Ø-mókk-* sketched above. That we are dealing of a wanderwort of Altaic origin thus seems conceivable, but its routes are not entirely clear.

## 5 *\*porkō-* and its many variants

The protoform *\*porkō-* (> Mlr *orc* [m. o-stem] 'young pig', Lat. *porcus* id., Lith. *pařsas* 'young pig; castrated male hog', OCS *prasę* 'young pig', Av. *pərəsō* id., PGmc. *\*farha-* 'pig' > e.g. OE *fearh*, OHG *farah*) is almost universally accepted as a major PIE 'pig'-word. Since Benveniste (1969) the meaning has been reconstructed as 'young pig' or 'piglet'<sup>25</sup>. However, as already noted by Hamp, the word is actually geographically confined to the Northwestern half of the Indo-European area. It is not found in Albanian, Greek, Armenian, Anatolian, Indo-Aryan, or Tocharian, and within Iranian, it is exclusively attested in the Northern fringes, geographically speaking. An Indo-Iranian preform is admittedly widely assumed as the basis for Fenno-Permian *\*poršas* or *\*porćas*

---

would be better. Kroonen further declares that the relationship of this word with the rhyming *\*rakka-* is unsolved. Since both Sw. *by-racka* 'mongrel dog' and *hundracka* 'cur', the loanword in Fi. *rakki* id., and MDu. *rekel* 'bad dog; male dog' (> Du. *rekel* 'villain') are clearly pejorative, I wonder if it is not simply derived from the verb *\*ragg/kōn-* 'to move to and fro, to stroll', cf. Sw. *racka* 'to roam, to wander about (used typically of dogs)'.

<sup>24</sup> Da. *gravhund*, lit. 'grave dog', cf. *grævling* 'badger', derived from *grav* 'grave'.

<sup>25</sup> As Hamp (189) says "The question is not one of domestication, but of cultural value".



‘pig’ (Benveniste 1949), but as some scholars (notably Napolskich 2002) have pointed out, the Balto-Fennic forms may just as well have been borrowed from a stage of Balto-Slavic, while the Eastern (Mordvinian and Permian) forms either suspiciously lack the Indo-Iranian ending or are aberrant in other ways. Outside Europe, the only attested forms are therefore Av. *pərəsō* and Khot. *pāsa* (Kurdish *purs* is a ghostword, cf. Hoffmann 1967: 35). These are both important culture languages of Central Asia, spoken close to Altaic and especially Turkic languages. As we shall see in a moment, there is a possibility that *\*porko-* is really a Central Asiatic culture-word.

First, it is important to note that the main justification for starting to look for a non-Indo-European source for *\*porko-* is the variation with which similar words with similar meanings occur in Europe, displaying a remarkably colorful variation of irregular correspondences: Alongside *\*porko-* it seems necessary to reconstruct a variant *\*porġo-* as the basis for PGmc. *\*farkīna-* ‘pig’ and CSL. *\*porzō* ‘boar; ram; bull’.<sup>26</sup> Furthermore, PGmc. *\*barga-* (OHG *barug*, OE *bearg*, Eng. *barrow*, OIc. *börgr*, MidDu. *barg*) would have to go back to a PIE form *\*b<sup>h</sup>orko-*, *\*b<sup>h</sup>orġo-* or *\*b<sup>h</sup>org<sup>h</sup>o-*.<sup>27</sup> If the Germanic protoform was actually *\*baruga-*, the stem *\*baru-* correspond to Slavic forms like Ru. *bórov* [m.] ‘hog, castrated boar’, SCr. *brâv* [m.] ‘sheep’, dial. ‘hog, castrated boar’. Finally, OHG *bēr*, OE *bār* ‘boar’ most likely go back to PGmc. *\*baira-*.<sup>28</sup> Pre-Proto-

<sup>26</sup> It is perhaps conceivable, but hardly provable, that *\*porzō* was reshaped from earlier *\*porsō* under the influence of Slavic *\*kōn-orzō* ‘(domesticated) boar’, lit. ‘with testicle’ (on this form, see Kretov 1994), and that PGmc. *\*farkīna-* can theoretically be interpreted as a diminutive *\*farhkīna-* of *\*farha-*. All things considered, I find it more economical and therefore more probable to assume that we are dealing with a true irregular variation, and that both forms go back to what would have been PIE *\*porġo-*.

<sup>27</sup> Not to be reconstructed *\*baruga-*, cf. Ball & Stiles 1983.

<sup>28</sup> *\*baiza-* is also possible. Kroonen (2013) regards *\*baira-* as more likely because the ON hypocoristic form *bassi* is more likely to have been derived from *\*bárr* < *\*baira-* than *\*beirr* < *\*baiza-*. However, *bassi* may just as well be a hypocoristic form of *baggi* ‘small and thick, compact animal’, cf. Da. *basse* ‘piglet’, but also ‘small, fat male horse, wether, dog, bull etc.’, ‘thick insect’, rather the same meanings as those of *baggi*. Polomé (1986) and Schrijver (1997), in the light of the irregular correspondence with MW *baed*, W *baedd*, OCorn. *bahet* ‘boar’, conclude that both the Germanic and the Celtic form are of the same non-Indo-European origin. However, if the PGmc. form was really *\*baira-* and especially if at the same time the Brythonic forms are from *\*badġo-* and not from *\*baġio-*, in my opinion they are not even sufficiently alike for us to regard them as the same word.

Germanic even had a *\*pōr-o-* (PGmc. *\*fōra-*), but this is clearly derived from *\*per-* ‘to give birth’, Gmc. *\*farzan-*.

therefore think that the irregularities of the Fenno-Permic forms are due to the fact that some of them were borrowed from R-Turkic, cf. the Chuvash form *porəš* ‘badger’ (via ‘piglet’? Note the exact meaning of *\*porkos*) with initial *p-* and without reminiscences of the suffix *-uk*, into the Eastern Fenno-Permic languages: Mordvin (Moksha) *puřć*, *purc*, Udmurt *parś*, *parəś*, *pars*, *paris*, Komi *pors*<sup>29</sup>, with their *\*-ś-* actually reflecting Turkic *-s-* (= Chuvash *-š*) and not an alleged PIE *\*-k̑-*. The Balto-Fennic languages (*\*porsas*) and perhaps Erzya Mordvin (*purcos*, *puřćis*, *pursuz*) instead borrowed their forms from a stage of Balto-Slavic, which comprised the still productive ending *-as*. The Balto-Slavic form, in turn, like the Germanic and Italo-Celtic forms, would ultimately have derived from older Turkic *\*borsuq*. Again, the exact directions are not clear, but a culture-word situation would account for the geographical distribution, the irregular variations of similar ‘pig’- and ‘badger’-words across Europe, aberrancies in Fenno-Ugric and perhaps even the exact meaning of *\*porkos* if the meaning went via ‘young of badger’ (Da. *gris*, also ‘pig’). Finally, the motivation for a borrowing from Altaic would be straightforward, namely the fact that badger is a common traditional dish among Turkish and Mongolian peoples.<sup>30</sup>

## 6 Conclusions

Etymological analyses that take Uralic material into account confirm Hamp’s (1987) claims that 1) most Celtic terms for pigs derive from a non-Indo-European language in Northern Europe, and that 2) the pig occupied a special cultural position among the speakers of this language. Three terms specifically can be traced directly back to the Balto-Fennic protolanguage (Late or Middle Proto-Fennic), and it seems possible to link this to extralinguistic data ranging from prehistoric archaeology to

<sup>29</sup> From Komi it has been borrowed into Khanty as V *porəs*, DN *purəš*, O *porəś*; and into Mansi as KU *pürs*, P *pōrəs*, So. *pūrəś*; Nenets Sj. *pors*; and from Khanty it has been borrowed further into Nenets O as *porəś*, Nj. *pořes*.

<sup>30</sup> At the conference leading to this publication, I argued that Turkic *\*borsuq* could have been borrowed to Gmc. as *\*barzúḡa-* yielding *\*bárruḡa-* that could then have been transferred to PCelt. *brokko-*, involving the same stress shift to the penultimate syllable *-bVr-úk-* as in *\*mokku-* < *\*e-mók-*. However, this is of course not possible if the PGmc. form was in fact *\*barga-* and not *\*baruga-*. See above on the role of PFC. *\*mäkrä*.

historical accounts. These conclusions have important implications both linguistically and culturally since they strengthen the hypothesis sketched in Hyllested (2010) that lexical exchange took place directly between the speakers of Proto-Fennic and Proto-Celtic, and that parts of the Celtic vocabulary have a Fennic origin. Other Northern European pig-terms are of Indo-European origin, but their exact reconstruction in some cases is only possible if Uralic material is analyzed on an equal footing. Studies in language contact within a given region can benefit from the involvement of all languages actually spoken in that region. It seems worthwhile to investigate whether other portions of the IE lexicon of Northern Europe that allegedly derive from long-gone substrata, are actually just loanwords from (different stages of) the adjacent language family, Uralic.

## References

- Ball, C.J.E. & Patrick V. Stiles, 1983: "The derivation of Old English *geolu* 'yellow', and the relative chronology of smoothing and back-mutation". – *Anglia* 101. 5–28.
- Balles, Irene, 2010: "Some new Celtic and other etymologies". – Giovanni Belluscio & Antonio Mendicino (eds.): *Scritti in honore di Eric Pratt Hamp per il suo 90 compleanno*. Cosenza: Università di Calabria. Pp. 15–20.
- Bammesberger, Alfred & Simas Karaliūnas, 1998: "Zur Frage nach der ethnischen Identität der Aisten". – Alfred Bammesberger (ed.): *Baltistik: Aufgaben und Methoden*. Heidelberg: Carl Winter. Pp. 39–47.
- Bellquist, J.B., 1993: " "Badger" in Indo-European". – *Journal of Indo-European Studies* 21: 321–346.
- Benveniste, Émile, 1949: "Le nom du porc". – *Bulletin de la Société Linguistique de Paris* 45: 74–91.
- Benveniste, Émile, 1969: *Le vocabulaire des institutions indo-européennes*. Paris.
- Blažek, Václav, 2010: "Balto-Slavic *ueprjo-* 'boar' ". – *Baltistica* XLV(I): 85–94.
- Campbell, Lyle, 1980: "The psychological and sociological reality of Finnish vowel harmony". – Robert Michael Vago (ed.): *Issues in vowel harmony: Proceedings of the Cuny Linguistics Conference on Vowel Harmony, May 14, 1977* [= *Studies in Language Companion Series* 6]. Amsterdam: John Benjamins. Pp. 245–270.
- Demiraj, Bardhyl, 1997: *Albanische Etymologien: Untersuchungen zum Albanischen Erbwortschatz* [= *Leiden Studies in Indo-European* 7]. Amsterdam & Atlanta (GA): Rodopi.
- Hamp, Eric P., 1986: "Culhwch, the Swine". – *Zeitschrift für Celtische Philologie* 41. 257–258.
- Hamp, Eric P., 1989: "North European Pigs and Phonology". – *Zeitschrift für Celtische Philologie* 43: 192–193.

- Hamp, Eric P., 1987: "The pig in ancient Northern Europe". – S.N. Skomal & E.C. Polomé (eds.): *Proto-Indo-European: The archaeology of a linguistic problem* [Fs. Gimbutas]. Washington, D.C.: Institute for the Study of Man. Pp. 185–190.
- Hoffmann, Karl, 1967: "Drei indogermanischen Tiernamen in einem Avesta-Fragment". – *Münchener Studien zur Sprachwissenschaft* 22: 29–38.
- Huld, Martin, 1984: *Albanian Etymologies*. Bloomington: Slavica.
- Hyllested, Adam, 2010: "The precursors of Celtic and Germanic". – Stephanie Jamison, Craig Melchert and Brent Vine (eds.): *Proceedings of the 21<sup>st</sup> annual UCLA Indo-European Conference*. Bremen: Ute Hempen. Pp. 107–128.
- Hyllested, Adam, forthcoming: "Baltic etymology as a key to Fennic ethnonyms". Ms.
- Hyllested, Adam & Bernd Gliwa, 2009: "Metatony in Lithuanian internal derivation". – Thomas Olander & Jenny Helena Larsson (eds.): *Stressing the past: Papers on Baltic and Slavic Accentology* [= *Studies in Slavic and General Linguistics* 35]. Amsterdam: Rodopi. Pp. 47–52.
- Häkkinen, Jaakko, 2007: *Kantaauralin murteutumisen vokaalivastaavuuksien valossa*. MA thesis from the University of Helsinki.
- Janhunen, Juha, 2009: "Proto-Uralic – what, where, when?". – *The Quasiquicentennial of the Finno-Ugric Society* [= *Mémoires de la Société Finno-Ougrienne* 258]. Helsinki: Suomalais-Ugrilainen Seura. Pp. 57–78.
- Kallio, Petri, 2007: "Kantasuomen konsonanttihistoriaa". – Jussi Ylikoski & Ante Aikio (eds.): *Sámit, sánit, sátnehámit: Riepmočála Pekka Sammallahtii miessemánu 21. beaivve 2007* [= *Mémoires de la Société Finno-Ougrienne* 253]. Helsinki: Suomalais-Ugrilainen Seura. Pp. 229–249.
- Kallio, Petri, 2012: "The non-initial vowel reductions from Proto-Uralic to Proto-Finnic". – Tiina Hyytiäinen, Lotta Jalava, Janne Saarikivi & Erika Sandman (eds.): *Per Urales ad Orientem: Iter polyphonicum multilingue. Festschrift tillägnad Juha Janhunen på hans sextioårsdag den 12 februari 2012* [= *Mémoires de la Société Finno-Ougrienne* 264]. Helsinki: Suomalais-Ugrilainen Seura. Pp. 163–175.
- Karulis, Konstantīns, 1992: *Latviešu etimoloģijas vārdnīca*. Riga: Avots.
- Katz, Hartmut, 2003: *Studien zu den älteren indoiranischen Lehnwörtern in den uralischen Sprachen*. Heidelberg: Winter.
- Kerkhof, Peter Alexander, forthc.: "Who brings home the \*bakkōn?". Ms.
- Khabtagaeva, Bayarma, 2009: *Mongolic elements in Tuvan*. Wiesbaden: Harrassowitz.
- Kretov, Aleksej A., 1994: "Baltijskie i slavjanskije prodolženija i.-e. \*orġh-". Stat'ja I: lit. *išarža* i prasl. *\*kǫnorzǫ*. – *Linguistica Baltica* 3: 197–206.
- Kristiansen, Kristian, 2009: *Proto-Indo-European Languages and Institutions: An Archaeological Approach*. – Marc Vander Linden & Karlene Jones-Bley (eds.): *Departure from the Homeland: Indo-Europeans and Archaeology*. Washington, DC: Institute for the Study of Man. Pp. 111–140.
- Kristiansen, Kristian & Thomas Larsson, 2005: *The Rise of Bronze Age Society: Travels, Transmissions, and Transformations*. Cambridge: Cambridge University Press.
- Kroonen, Guus, 2011: "Neuniederländisch zeug 'Sau' und *spugen* 'speien': zwei Fällen von paradigmatischem Wechsel der Westgermanischen Verschärfung?". – *Amsterdamer Beiträge zur älteren Germanistik* 67 [Fs. Quak]: 145–157.
- Kroonen, Guus, 2013: *Etymological Dictionary of Proto-Germanic*. Leiden: Brill.

- Kümmel, Martin, 2004: "Ungeklärtes \*-u- neben Liquida in Germanischen Nomina". – A.Hyllested, A.R. Jørgensen, J.H.Larsson & Th. Olander (eds.): *Per Aspera ad Asteriscos* [Fs. Rasmussen]. Innsbruck: IBS. Pp. 291–303.
- Larsson, Åsa M., 2003: "Uniting Strategies. Material Culture in Eastern Middle Sweden at the End of the Middle Neolithic". – Åsa M. Larsson & Ludvig Papmehl-Dufay (eds.): *Uniting Sea II: Stone Age Societies in the Baltic Sea Region*. Uppsala: Institutionen för Arkeologi och Historia. Pp. 132–149.
- Lubotsky, Alexander, 1994: "Avestan θβərəstar- and the Indo-European root √turk-". – *Die Sprache* 36/1: 94–102.
- Mallory, James P. & Douglas Q. Adams, 2006: *The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World*. Oxford: Oxford University Press.
- Malmström & al., 2009: "Ancient DNA reveals Lack of Continuity between Neolithic Hunter-gatherers and contemporary Scandinavians". – *Current Biology* 19,1–5. doi:10.1016/j.cub.2009.09.017
- Matasović, Ranko, 2009: *Etymological Dictionary of Proto-Celtic*. Leiden: Brill.
- McCone, Kim, 1992: "OIr. *torc*, Av. θβərəsō < PIE \*tworkós "(cutter,) boar" ". – *Münchener Studien zur Sprachwissenschaft* 53: 99–100.
- McCone, Kim, 1993: "The etymology of Old Irish *torc* 'boar' ". – Bela Brogyanyi & Reiner Lipp (eds.): *Comparative-historical linguistics: Indo-European and Finno-Ugric. Papers in honor of Oswald Szemerényi III*. Pp. 291–292.
- Napolskich, Wladimir, 2002: "Zu den ältesten Beziehungen zwischen Finno-Ugriern und zentraleuropäischen Indogermanen". – R. Blokland & C. Haselblatt (eds.): *Finno-ugrians and Indo-Europeans. Linguistic and literary contacts*. Maastricht: Shaker. Pp. 265–271.
- Olsen, Birgit Anette, 2012: "Contributions to Armenian etymology". – Stephanie Jamison, Craig Melchert & Brent Vine (eds.): *Proceedings of the 23<sup>rd</sup> annual UCLA Indo-European Conference*. Bremen: Ute Hempen. Pp. 113–124.
- Orel, Vladimir, 1998: *Albanian Etymological Dictionary*. Leiden, Boston & Cologne: Brill.
- Polomé, Edgar C., 1986: "The non-Indo-European component of the Germanic lexicon". – Annemarie Etter (ed.): *O-o-pe-ro-si. Festschrift für Ernst Risch zum 75. Geburtstag*. Berlin & New York (NY): Walter de Gruyter. Pp. 661–672.
- Quak, Arend, 1983: "Zu den salfränkischen Tierbezeichnungen". – *Amsterdamer Beiträge zur älteren Germanistik* 19: 7–66.
- Ritter, Ralf-Peter, 1975: "Slavisch \*ežyo- 'Dachs' ". – *Die Sprache* 21: 191–194.
- Ritter, Ralf-Peter, 2005: "Zur Herkunft von fi. *mäyrä* 'Dachs' ". – Cornelius Haselblatt, E. Koponen & Anna Widmer (eds.): *Lihkkun lehkos! Beiträge zur Fenougristik aus Anlass der sechzigsten Geburtstags von Hans-Hermann Bartens* [= *Veröffentlichungen der Societas Uralo-Altaica* 65]. Wiesbaden: Harrassowitz.
- Rowley-Conwy, Peter & Jan Storå, 1997: "Pitted Ware seals and pigs from Ajvide, Gotland: Methods of study and first results". – G. Burenhult (ed.): *Remote Sensing*, vol. 1. *Teses and Papers in North-European Archaeology*. Stockholm: Stockholm University. Pp. 113–125.
- Sainéan, Lazare, 1905: *La création métaphorique en français et en roman: images tirées du monde des animaux domestiques II: Le chien et le porc avec des appendices sur le loup, le renard et les batraciens* (Beihefte zur Zeitschrift für romanische philologie. 1., 10. hft.). Halle an der Saale: Max Niemeyer.

- Schindler, Jochem, 1966: "Germanisch \**selxaz* 'Seehund' ". – *Die Sprache* 12: 65–66.
- Schrijver, Peter, 1997: "Animal, vegetable and mineral: some Western European substratum words". – Alexander Lubotsky (ed.): *Sound law and analogy: papers in honour of R.S.P. Beekes*. Amsterdam: Rodopi. Pp. 293–316.
- Smoczyński, Wojciech, 2003: *Studia Balto-Słowanskie* 2. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
- Testen, David. 1999: "Stem-final \*-*kk*- in Celtic terms for 'pig' ". – *Ériu* L: 161–164.
- Wagner, Heinrich, 1954: "Varia 4. Ir. *Broc*, kymr. *Broch* 'Dachs' ". – *Zeitschrift für Celtische Philologie* 24: 92–93
- WGS = August Fick, Hjalmar Falk & Alfred Torp, 1909: *Wörterbuch der Indogermanischen Sprachen: Dritter Teil: Wortschatz der Germanischen Spracheinheit*. Göttingen.



## The Other Horse: Germanic Cognates of *caballus*?

### Abstract

Simon's (2005) reconstruction of a second PIE term for 'horse' as the source of Lat. *cabō*, *caballus*, Iranian *\*kaba-*, OCS *kobyła* and Fi. *hepo* (via Germanic) is accepted, but its shape must have been *\*keb<sup>h</sup>-*, not *\*keb-*, and the PGmc. form was *\*hebō*, not *\*hepa-*. The Germanic evidence is not restricted to loanwords in Finnish; in this article, Da. *hoppe* 'mare' and its Germanic relatives are interpreted as the old oblique form of a PGmc. n-stem paradigm *\*hebō* ~ *\*huppaz* whose nominative was transferred into Northern Balto-Fennic as *hepo* while the Southern languages borrowed *\*huppaz* (perhaps even in a Pre-Germanic shape with *\*-p-n-* before Kluge's Law operated). This solution provides the first serious explanation for the irregular variation within Balto-Fennic, notably the difference between Fi. *hepo* and Est. *hobune*, *hopene*.

### 1 Reconstructing PIE *\*keb<sup>h</sup>-* 'the slow horse'

Alongside the famous PIE term *\*h<sub>1</sub>éḱuos* 'horse' (most recently treated by Huld 2004, de Vaan 2009, Blažek 2010), Simon (2005) has reconstructed another Indo-European word for this animal, PIE *\*keb-*, on the basis of the following forms:

- a) OCS *koby-la* 'mare' (< *\*kob-ōn*; borrowed into MHG dial. as *kobel* 'jade, hack, nag, crowbait', and, as a back-formation, Swabian *kob* 'id.')

---

<sup>1</sup> Mod.Eng. *cob* 'thickset horse' might be unrelated if simply identical in origin to the numerous other meanings of Eng. *cob*, most of which can be united by the meaning 'plump or roundish object, animal or person'.



- b) late gloss Lat. *cabō* ‘castrated horse’ (< *\*keb-ōn*);
- c) Iranian *\*kaba-*, *\*kabala-* (borrowed into Greek as καβάλλης ‘working horse’ [Hes.] and into Latin as *caballus*) > Khotan Saka *kabā* ‘horse’ (borrowed into Karakhanidic Middle Turkic [11<sup>th</sup> c. AD] as *kevel* (at) ‘quick horse’), Mod.Pers. *kawal* ‘mule’; and
- d) PGmc. *\*hepa-*, reconstructed solely to account for supposed loanwords in Balto-Fennic languages, e.g. Fi. *hepo*, *hevonen* ‘horse’

Simon gives convincing arguments for leaving out Slavic *\*kon’b* and *\*komonb* (pace Snoj 2003) and for rejecting alternative loan directions such as Greek into Persian. On the whole, the proposal looks quite acceptable, despite Indo-Europeanists’ usual reluctance to accept inherited lexemes with the rare unaspirated *\*b* for PIE.

On closer inspection, it turns out that it is unnecessary to reconstruct this unaspirated *\*-b-* since the only element that precludes its aspirated counterpart *\*-bʰ-* is exactly the PGmc. *\*-p-* which Simon supposes to only be reflected in Balto-Fennic loans. Since Balto-Fennic did not possess voiced stops<sup>2</sup>, the language would be unable to show evidence of original voicing, so PGmc. *\*-b-* < PIE *\*-bʰ-* is just as possible in this instance. In fact, the lack of Winter’s Law in Slavic (not †*kabyła*) unambiguously points to an aspirate, if *kobyła* is inherited at all (see Blažek 2010 on the East Iranian animal-term suffix *\*-ūla-*; Gołąb 1985).

## 2 The problem of Balto-Fennic variation

A serious obstacle not addressed by Simon is that only the forms occurring in the Northern Balto-Fennic languages (Finnish, Karelian, Lude and Veps) differ from each other in regular ways. Most forms in the Southern languages (Votic, Estonian, South Estonian, and Livonian) exhibit variation which is both internally irregular and deviates from the Northern forms. The biggest issue is the vocalism with *\*-o-*, but the variation in the suffixes between *\*-u-* and *\*-e-* and between *\*-p-* and *\*-pp-* is not unproblematic either. Next to *hebu*, Estonian variants include

<sup>2</sup> In Modern Finnish, voiced stops occur in very recent loanwords, and *\*-d-* occurs as the weak form of *\*-t-* in consonant gradation, but in Balto-Fennic times it is reconstructed as a voiced dental fricative.

*hobune, hobu, hobene* (<b> = /p/), even *hopen* with geminate /p:/ (rendered in writing as <p>) and Votic has *opõn* next to *õpõn*:<sup>3</sup>

Northern Balto-Fennic:

Finnish	Karelian	Lude	Veps
<i>hepo</i>	<i>hebo</i>	<i>hepo</i>	<i>hebo, h'ebo, h'öbo</i>
<i>hevonen</i>	<i>heboñe</i>	<i>heboiñe</i>	S dial. <i>hebõñe</i>
Ingr.dial. <i>heppoin</i>	Olonets dial. <i>heboine</i>		

→ Kukkuzi Votic  
*hepoina*

Southern Balto-Fennic:

Votic	Estonian	Livonian
-	dial. <i>hebu</i>	-
-	<i>hobu, hobo</i>	<i>õ'bbi, y'bbi, i'bbi, ü'bbi</i>
<i>opõn, õpõn</i>	<i>hobune, hobene, hopen</i>	-

The variation in both vocalism and consonantism between different Balto-Fennic forms has been close to ignored in virtually all etymological proposals for *hepo*. To mention some of the most recent proposals, Liukkonen (1999) argues for an origin in Baltic *\*ešva-* with metathesis, cf. Lith. *ašvienis* ‘working-horse’ that also has a nasal suffix; and Kortlandt (1997) surmises a loan directly from (some stage of) PIE *\*h<sub>1</sub>ékʷos* with the laryngeal preserved<sup>4</sup>. LÄGLOS (I: 95-96) does mention the

<sup>3</sup> Est., Vot. and Liv. *õ* represents an unrounded mid-high back-vowel which most often originates from BF *\*e*.

<sup>4</sup> It is tempting to suggest as an alternative Iranian etymology of the word, involving completely different elements, cf. Chin. *chībō* < Middle Chin. *\*ts'iat puat*, borrowed from Sogdian *čʾrθpād-* (*čyrδpδ*) ‘horse used in the valleys’, a formation identical to Lat. *quadru-pēs* ‘four-legged (animal)’ (Yoshida 2009). Uralic, Fenno-Ugric, Fenno-Permian and Fenno-Saami *\*č* become *\*h-* in Balto-Fennic. However, neither such a proposal would solve the issue of variation in Balto-Fennic material; even if one disregarded the suffixes and argued that the alternation between *\*-pp-* and *\*-p-* (~ weak grade *\*-β-*) could be due to different renderings of the Iranian cluster *\*-rθp-* the occasional *o*-vocalism would remain enigmatic.

Southern forms, but fails to account for their occurrence. The possibility of a Germanic origin (from < \*ehwa-), again with metathesis, is left open, but the entry ends with the judgement “Kaum ein germ. Lehnw.”.

### 3 Danish *hoppe* and its closest relatives

The earliest attestation of Da. *hoppe* ‘mare’ as a common noun is from 1621, but it is known much earlier from a place-name *Hoptrup* near Haderslev in S Jutland, 1287-1307 *Hoptorp*, 1289 *Hoppetarp*, 1421 *Hoptorpe*. Nothing points to an original meaning ‘mare’: The names correspond structurally to other towns in S Jutland: *Hostrup*, attested 1298 as *Horstrop* (near Tønder), ca. 1325 *Horsthorp* (near Esbjerg), 14th. c. *Hors-torp* (near Vejle) where the first member is *hors* ‘horse (obs.)’ (Jørgensen 1994). We also know that *hoppedreng* (obs., attested 1653) simply meant ‘groom, stableboy’ and did not refer to mares in particular<sup>5</sup>. Its Swedish counterpart is a false friend *hoppa* ‘old meagre horse’ (dial., att. 1683), and the Nordic word has been borrowed into OE as *hobin* > ME *hobyn* ‘nag’ > Eng. *hobby*. The semantic development ‘horse’ → ‘mare’ is quite common, cf. e.g. PGmc. \**marha-* > Older Da. *mar* ‘mare’ (alongside *mær* from the original fem. \**marhjōn-*) and PGmc. \**hursa-* ‘horse’ > Bornholm dial. of Da. *horse* ‘mare’.

On the basis of these forms and LG dial. *huppe*, we can reconstruct a PGmc. \**huppaz*. The standard Danish and Swedish handbooks (e.g. Svenska Akademiens Ordbok, SAOB) otherwise assume, often with some hesitation, that *hoppe* is an inner-Scandinavian derivative of the word *hoppe* ‘to hop, jump’ < \**huppōn-*. However, a homonymous verb existed in Germanic with the meaning ‘move backwards, retreat (especially of horses)’ and for the latter non-iterative variant \**happōn-*, and judging by the Swedish and English semantics, this is a more obvious connection (which however does not work completely, see below). So far, on the basis of these forms and LG dial. *huppe*, we can reconstruct a PGmc. \**huppaz* ‘slow horse’.

<sup>5</sup> The surname *Hoppe*, known from 1410 onwards, according to Gammeldansk Ordbog’s note collection ([www.gammeldanskordbog.dk](http://www.gammeldanskordbog.dk)) may instead be from *hoppe* in the meaning ‘shrimp’, derived from the verb *hoppe* ‘to hop’ (cf. also Mod.Da. *græshoppe* ‘grasshopper’), or from MLG *hoppe* ‘hop (the crop)’.

#### 4 A Germanic ablauting n-stem

I suggest that Proto-Germanic possessed an n-stem *\*hebō* ‘horse’ with gen. *\*huppaz* of the type with vocalism alternating between *-e-* and *\*-u-* described by Kroonen (2011). Although none of the items listed by Kroonen constitute an exact parallel to such a paradigm, i.e. no item with an alternation between nom. *\*-eD-* ~ gen. *\*-uTT-*<sup>6</sup>, one example of *\*-eT-* ~ *\*-uTT-* does occur (*\*wekō* ~ *\*wukkaz* ‘wick’) as does *\*-eD-* ~ *\*-uRTT-* as well as numerous other similar types like *\*-euD-* ~ *-uTT-* and *\*-aD-* ~ *\*-uTT-*. The system thus leaves the possibility open of the existence of a subtype *\*-eD-* ~ *\*-uTT-*.

If the Germanic form of the horse-word suggested by Simon (2005) was actually *\*hebō* ~ *\*huppaz* originating from *\*keb<sup>h</sup>-ōn-* ~ *\*k(ə)b<sup>h</sup>-nó-s*, the *-u-* at first glance would have to have arisen from an analogical Germanic use of *\*-u-* in zero-grade surroundings because no sonorant or *\*-u-* was ever present to trigger the Gmc. *-u-* vocalism in the weak stems on regular grounds (as in PIE *\*greb<sup>h</sup>-on-* ~ *\*grb<sup>h</sup>-nó-s* > PGmc. *\*krebō* ~ *kurppaz* ‘basket’). However, there is also the possibility that the ablaut of the iterative verbs was exported to the nominal level as Kroonen (2011: 211–212) suggests for parallel cases like *\*dabō-* ~ *\*duppaz* ‘puddle’ next to the verb *\*duppōn-*. The vowel *\*-u-* became productive as a zero-grade marker in Germanic iteratives even derived from strong verbs (cf. Kroonen 2012: 193). And it seems reasonable to assume that the iterative verb *\*hupp/bōn-*<sup>7</sup> and its non-iterative counterpart *\*happōn-*<sup>8</sup> ‘to move backwards, retreat (especially of horses)’ influenced the paradigm of *\*hebō*, especially if Simon is right that this word originally denoted the “slow” horse.

<sup>6</sup> Kroonen (2011: 335–351) rejects an original paradigmatic alternation *\*kredō-* (OHG *chreta*, MHG *krete*) ~ *\*kruttaz* (OHG *chrota*, MLG *kröde*) for ‘toad’, explaining the Upper German forms with <e> as renderings of umlauted forms with *-ö-* corresponding to NHG *Kröte*.

<sup>7</sup> Cf. ON *hopa* ‘fall back’, Icel. *hopa* ‘turn back, retreat’, Far. *hopa* ‘draw back, recede, retreat’, Nw. *hop(p)a*, *hobba* ‘retreat, drive backwards (esp. of horses)’, Da. dial. *hoppe sig* ‘move backwards, drive backwards’ (Kroonen 2013: 257).

<sup>8</sup> Cf. Nw. *haba*, *habba* ‘retreat, drive backwards (esp. of horses)’, Sw. dial. *habba* ‘to turn back, drive backwards’ (Kroonen 2013: 257).

## 5 From Germanic to Balto-Fennic

If the Balto-Fennic words for ‘horse’ are borrowed from such a Proto-Germanic *-n*-stem, the vocalic and consonantal variation among Balto-Fennic forms can be explained by the Germanic paradigmatic alternations. The Northern Balto-Fennic form *\*hepo* and Est. dial. *hebu* are then simply a rendering of the old Germanic nominative, while Southern Balto-Fennic borrowed oblique forms with new zero-grade, perhaps with the *-n*- of the stem still retained and reinterpreted as domestic derivational suffixes:

*\*hebō* → (loan) Fi. *hepo*, Est. *hebu*  
                   > †  
*\*huppaz* < *\*hup-na-* → (loan) Est. *hobune* /*hopune*/, *hopene* /*hop:ene*/ etc.  
                   > Da. *hoppe*, Sw. *Hoppa* (Scand. → OE *hobin*, ME *hobyn*, *hobin*,  
                   Eng. *hobby*), LG dial. *Huppe*

Both forms were easily incorporatable into the Balto-Fennic system because BF already possessed the nominal suffixes *\*-o* and *\*(i)nen*. It happens quite often that originally borrowed strings come to be interpreted as native suffixes (cf. e.g. BF *\*hom-eh* ‘fungus’ < *\*šomeš* ← Gmc. *\*swambaz*).

## 6 Concluding remarks

It cannot surprise us that yet another term for ‘horse’ turns out to be a culture-word as the invention of riding spread with human migrations, expansions and trade. The exact directions of culture-words are notoriously difficult to trace, not least because they often involve transmission via unattested languages, and their etymologies therefore often remain disputed. From a purely Indo-European perspective, Simon’s proposal might be considered as plausible as any other. It is his inclusion of borrowings into non-Indo-European languages that suddenly made the evidence for an inherited PIE word worth considering. Correspondingly, in this article, a closer look at the Balto-Fennic material led us to revise his Germanic etymology and ultimately the Indo-European one, and consulting the latest research in Proto-Germanic morphophonology made it possible for us to explain the otherwise enigmatic differences between Balto-Fennic forms.

Simon's observations already shed new light on the motivation of denoting the horse 'the quick one': because if *\*keb<sup>h</sup>-* was 'the slow horse' then *\*h<sub>1</sub>ékʷos* is not simply 'the swift one' understood as swift runners compared to other animals, but rather 'the swift (kind of) horse', i.e. the kind used for riding, as opposed to the slower working-horse.<sup>9</sup>

## References

- Blažek, Václav, 2010: "On 'horse' in Slavic". – Stephanie W. Jamison, H. Craig Melchert & Brent Vine (eds.): *Proceedings of the 21<sup>st</sup> Annual UCLA Indo-European Conference*. Bremen. Pp. 13-25.
- Gołąb, Zbigniew, 1985: "Slavic *komonb* and *kon'b* 'equus': An Attempt at Etymology against the Background of the History of Domestication". *Journal of Indo-European Studies* 13: 3-4: 415-443.
- Huld, Martin, 2004: "An Albanian reflex of Proto-Indo-European *\*E<sub>1</sub>ékʷos* 'horse'". – Karlene Jones-Bley, Martin E. Huld, Angela Della Volpe & Miriam Robbins Dexter (eds.): *Proceedings of the Fifteenth Annual UCLA Indo-European Conference*. Washington D.C. Pp. 186-195.
- Jørgensen, Bent, 1994 [1st ed. 1981-1983]. *Stednavneordbog*. Copenhagen.
- Kortlandt, Frederik, 1997: "Labials, Velars, and Labiovelars in Germanic". – *NOWELE* 30: 45-50.
- Kroonen, Guus, 2010: *The Proto-Germanic n-stems: A study in diachronic morphophonology* [= *Leiden Studies in Indo-European* 18]. Amsterdam.
- Kroonen, Guus, 2012: "Reflections on the o/zero-Ablaut in the Germanic Iterative Verbs". – H. Craig Melchert (ed.): *The Indo-European Verb: Proceedings of the Conference of the Society for Indo-European Studies, Los Angeles 13-15 September 2010*. Wiesbaden: Reichert. Pp. 191-200.
- Kroonen, Guus, 2013: *Etymological Dictionary of Proto-Germanic*. Leiden: Brill.
- LÄGLOS = A.D. Kylstra, Sirkka-Liisa Hahmo, Tette Hofstra & Osmo Nikkilä: *Lexikon der Älteren Germanischen Lehnwörtern in den Ostseefinnischen Sprachen I-III*. Amsterdam / New York 1991-2012.
- Liukkonen, Kari, 1999: *Baltisches im Finnischen*. Helsinki.
- Simon, Zsolt, 2005: "Die Etymologie von *caballus*". – Gualtiero Calboli (ed.): *Latina Lingua! Proceedings of the Twelfth International Colloquium on Latin Linguistics (Bologna, 9-14 June 2003)*. Roma. Pp. 405-416.
- Snoj, Marko, 2003: *Slovenski etimološki slovar*. Ljubljana.
- de Vaan, Michiel, 2009: "The derivational history of Greek ἵππος and ἵππεύς". – *Journal of Indo-European Studies* 37, 1/2: 198-213.
- Yoshida, Yutaka, 2009: "Sogdian". – Gernot Windfuhr (ed.): *The Iranian Languages*. London / New York: Routledge. Pp. 279-335.

<sup>9</sup> PIE *\*kōpHo-* 'hoof' (PGmc. *\*hōfaz*), despite the similarity, appears unrelated to *\*keb<sup>h</sup>-* even though ON *hófr* means both 'hoof' and 'horse'; this probably just reflects a *pars pro toto* semantic extension of the type ON *horn* 'horn', but also 'ox'.



# Balto-Fennic Loanwords in Proto-Germanic<sup>1</sup>

## Abstract

This article consists of four entries treating what is argued to be Proto-Balto-Fennic (Middle and Late Proto-Fennic) loanwords into Proto-Germanic. Many of the candidates for loanwords begin with *h-* – this feature makes them more easily detectable in cases where Balto-Fennic *h-* have cognates elsewhere in Uralic beginning with *š-* or *č-*.

## 1 Four entries

- 1.1 PGMc. \**HAMARA*- ‘HAMMER’ ← BF \**HAMARA* ‘BACK OF AN AXE’

PGmc. \**hamara-* m. (> ON *hamarr*, OE *hamor*, *hamer*, OFris. *hamer*, *homer*) has traditionally been viewed as inherited from PIE \**h<sub>2</sub>ék-mon-* or \**ák-mon-* ‘(sharp) stone’, yielding Slavic \**kamy*/\**kamen-* ‘stone’, Lith. *ašmuō* and *akmuō*, Gk. ἄκμων ‘anvil’, Skt. *áśman-* and Av. *asman-* ‘stone, sky’ (via ‘vault’). Just like PGmc. \**wajju-* ‘wall’ as a derivative of PIE \**ueiH-* ‘to wind, plait’ tells us that there was once a time when walls consisted of plaited branches, quite independently of any archaeological evidence, we also think because of the standard etymology that we are able to tell that hammers in Pre-Proto-Germanic times were made of stone.

However, it is by no means certain that PGmc. \**hamara-* is really cognate with the other forms mentioned. Slavic \**kamy*/\**kamen-* is mysteriously distorted. For those who reconstruct a laryngeal in the PIE word it looks like a late metathesis: \**h<sub>2</sub>ák-mon-* > PBSl. \**Hák-men-* >

---

<sup>1</sup> This paper was presented at the symposium “Germanic, Romance and Slavic in the Early Middle Ages” at the University of Leiden, 29 November 2012.



\*káH-men. However, the vowel PGmc. \*hamara- is short, which rules out the metathetic explanation that works for Slavic. For those of us who prefer an initial a-vowel, a metathesis does not even suffice because Slavic -a- reflects a long vowel.

Furthermore, PGmc. \*hamara- is a masculine thematic stem, and animate \*-men-stems are normally not heteroclitic. If it is Indo-European, it could be either a secondary thematicization of an old \*-mer/\*-men heteroclit, or the thematic suffix \*-ero- added to a root ending in \*-m-. The existence of a neuter \*ák-mŋ/-m(e)n- underlying the masculine 'stone'-word is possible but there seems to be no surviving reflexes of such a form. Indic derivatives such as the adjective *aśmará-* 'stony' can just be interpreted as \*ák-mŋ- + suffixal \*-ro-.

ON *hamarr* meant: (1) hammer, (2) back of an axe, (3) crag, precipice (rather than just "stone"); in compounds it could be used of rocks, e.g. *berghamarr* 'rocky precipice', *hamarrifa* 'rift in a crag' (Zoëga 1910). The latter meaning, however, is more likely figurative than primary. Correspondingly, in the Da. place-name *Hammer Odde*, *Hammerknuden* (on Bornholm, 1539 *Hammar*), *Hammer Bakker* (1503 *Hammer*) the name refers to the hammerhead-shaped crag of granite – again, the meaning 'hammer(head)' is primary, its use of a rock is a figurative description of a steep rugged mass of rock projecting outwards and upwards. It mostly occurs in coastal areas and never seems to be used just of rocks or stones in general that are not protruding or hammer-shaped. This use is confirmed by a common noun *hammer* '(steep) crag' in Danish dialects. There is thus no particular need to reconstruct the meaning 'stone' or 'rock' for PGmc. \*hamara-.<sup>2</sup>

The first and second meaning of the Old Norse word match perfectly that of Balto-Fennic \*hamara (> Fi. *hamara* 'back of an axe', Est. *hamar*, *hammar* 'back of a knife'). In my opinion, PGmc. \*hamara- has not been borrowed into Proto-Balto-Fennic as is otherwise assumed, but is simply a borrowing the other way around<sup>3</sup>. The most important reason is that the word must be inherited in Balto-Fennic from at least the Feno-Volgaic stage. It can be connected to Saami \**symērē* 'back/pole of an

<sup>2</sup> Scholars in general seem to regard the meaning 'rock' as primary in Germanic, but that may simply be circular reasoning – i.e. because they suppose from the outset that it is connected to PIE \**h<sub>2</sub>ek-men-*.

<sup>3</sup> Or at least into NW Germanic since no reflexes are known from Gothic. However, it must have entered the NW Germanic area already in the 2<sup>nd</sup> or 3<sup>rd</sup> c. AD because it occurs in the personal name *Chamarus* in the Zülpich-Enzen dedication: Lat.-Gmc. *Matronis M(arcus) Chamari f(ilius) et Allo* 'To the mother goddesses, Marcus, son of Chamarus, and Allo'.

axe, back of knife' (> N Saami *šibmar*, Lule Saami *šjimēr*, Skolt Saami *šammer*) and Mordvin *\*šuvV* 'back of a knife' (Erzya *čov*, *čovone*, Moksha *šov*). The Saami forms at first glance appears irregular with its *\*šim-* for expected *\*sam-*, but Western Saami *\*ši-* (< Proto-Saami *\*sī-*) here must actually be the regular outcome of Fenno-Ugric/-Permic/-Volgaic *\*ša-* in the position before nasal (the only other example being *\*šama* 'apparition, shape' mentioned below). While Uralic and Fenno-Volgaic *\*š-* indeed normally yields Saami *\*s-*, the actually attested forms, which are comparatively few, reveal that when followed by the vowel *-a-* the sequence yields Saami *\*ši-* before *\*-m-* (probably: any nasal, but there are no examples with *\*-n-* and *\*-ŋ-*); in fact there seems to be only one alternative surrounding attested which is *\*su-* before a labial stop (N Saami *suhpi* 'asp' ~ Fi. *haapa*). When followed by other vowels, both these vowels and *\*š-* behave as expected (e.g. N Saami *savvi-* 'heal a wound' < FV *\*šeŋä* and N Saami *soarvi* 'dead pine-tree < PU *\*šorwa-* 'to dry (out)'). *\*ča-* in front of a nasal is not affected either, cf. N Saami *čuoži* 'membrane, fleshy fibres on the inner side of the skin' (*\*-m-* regularly goes to Ø via *\*-w-* in these surroundings, cf. *muošmi* < Pre-PGmc. *\*mams-ma-*; see elsewhere in this publication)<sup>4</sup>.

There is a way to get around it all: The Fenno-Volgaic form *\*šamara* might, in turn, originally have been borrowed from some stage of Balto-Slavic. This would have been a language with regular satem reflex but the same metathesis or ablaut form as Slavic *kamy*, Slovencian *kamor*. Thus, a possibility is left open that PGmc. *\*hamara-* could be connected to *\*ak-men-* after all, however not as a direct reflex of its protoform.

## 1.2 PGMC. *\*HAMA-* 'SHAPE' ← BF *\*HAAMO* 'SHAPE'

PGmc. *\*hama-* or *\*haman-* 'shape, physical form' (> ON *hamr* 'skin of animal; guardian spirit, tutelary spirit', Da., Nw. *ham* 'skin of animal; apparition, ghost', OS *hamo* 'covering', OE *hama*, *homa* 'id.', OHG *hamo* 'skin; covering; fishing-net')<sup>5</sup> can likewise be a borrowing from Pro-

<sup>4</sup> Initial *\*š-* in Samic is basically regarded as a loanword phoneme, consistently indicating loans from Balto-Fennic, Baltic and Late Scandinavian. The inherited postalveolar fricatives turned into dental-alveolars in Proto-Saami (PU *\*č*, *\*š* → PS *\*c*, *\*s*). The conditioning presented here might indicate that it rather deserves to be characterized as a "marginal" phoneme, most common in loanwords but also occurring elsewhere.

<sup>5</sup> Kroonen (2013) alongside the *n*-stem adds a *ja*-stem *\*hamja-* on the basis of Mod.Icel. *hem* and Elfdalian *iem* 'fish skin'. Bjorvand & Lindeman (2000: 342)

to-Balto-Fennic. For BF we must reconstruct *\*haamo* ‘shape; apparition’ on the basis of Fi. *hahmo*, *haamu* ‘apparition, facial characteristics, looks; ghost; shadow’, Karelian *hoamu* ‘shape, contour; apparition, ghost’, Votic *āmo* ‘id.’ The medial *-h-* in standard Fi. *hahmo* is secondary as indicated by the shape of its cognates. The medial *-h-* in Finnish can either have arisen by influence of the initial /h/, as also in *huhmar* ~ *huumar* ‘mortar’; via metathesis from a derivative *\*haamo-h-*; or less likely, by transmission via an extinct Sami language, as in *kahlata* ~ *kaalata* ‘to wade’ ← Proto-Saami *\*kālē-* < PU *\*kälä-*).

This form, in turn, goes back to Fenno-Volgaic *\*šama* which also yields Lule Saami *šjipmō* ‘similarity’ and Mordvin M *šama*, E *čama* ‘face’. The Saami form at first glance appears irregular with its *\*sjim-* for expected *†suom-*, but as mentioned before, Saami *\*ši-* here must actually be the regular Saami outcome of Fenno-Ugric/-Permic/-Volgaic *\*ša-* in the position before nasal (the only other example being *\*šama-ra* mentioned above).

The aforementioned *čuoži* ‘membrane, fleshy fibres on the inner side of the skin’ is a reflex of a similar Fenno-Ugric root which exhibits almost the same semantics as *\*šama*: FU *\*čamčV* ‘skin (of animal), membrane’ is reconstructed on the basis of forms from all Saami languages compared with Khanty (Ostyak) *čunč* ‘skin, hide’. The reconstruction of Proto-Saami *\*-m-* in this root is based exclusively on Kildin Sami *tsūmts*, while all other Saami languages point to a protoform *\*cōncē*. However, that a reflex of a sequence *\*-mD-* (i.e. *-m-* + dental) is preserved only in a single daughter-language and assimilated everywhere else is typologically common because (also sporadic) dental assimilation of *-m-* before a dental is in itself very common<sup>6</sup>. Crucially, a reconstruction *\*cōncē* would leave the Kildin form unexplained, while *\*cōmcē* accounts for all of them (by preservation, in the case of Kildin Saami, or assimilation). Therefore I will reconstruct Proto-Saami *\*cōmcē*. It should be noted that even if the Proto-Saami form was *\*cōncē* (which I find unlikely), this form, in turn, could of course just as easily as the individual Saami forms be the product of sporadic assimilation of an original cluster with *\*-m-*.

---

reconstruct a PGmc. *\*hami-*, which could form the basis of *\*hamja-*, next to *\*haman-*, while Orel (2003: 158) only sees a thematic stem *\*hama-* m.

<sup>6</sup> The situation is parallel to Indo-European where usually only Baltic preserves *\*-mD-* whereas all other language groups assimilate (not necessarily by a rule-governed development). Note also the assimilation of Uralic *\*-mt-* > Fenno-Ugric *\*-nt-* in the ordinal suffix (and in lexemes as well).

The Mordvin forms might just as well < \*čama. Since Balto-Fennic \*h- is the result of PU \*č-, too, we might as well assert \*čama as the source. Both \*čama and \*čamčV are reconstructed for stages later than Proto-Uralic, and their variation may be due to their identity as Eurasian culture-words, cf. e.g. Written Mongolian čamča ‘shirt’ ~ Persian jāma ‘apparel’.

1.3 PGMC. \*HALBA- ‘HALF’ ← BF \*HALPA, HALBA- ‘CHEAP, REDUCED’

Among the classical, more or less successful, attempts to etymologize the PGmc. adjective \*halba- ‘half’ is the reconstruction of a derivative \*kol-b<sup>h</sup>o- with the adjectival suffix \*-b<sup>h</sup>o- added to the root of Lith. šalis ‘side; strip of land; coast; direction’<sup>7</sup>. The original meaning would then have been ‘belonging to one side, having to do with one side’, and such a development finds a semantic parallel in Slavic (OCS polъ means both ‘side’ and ‘(a) half’. The connection with šalis is, however, uncertain; it is not favored by Orel (2003: 154) who labels it “difficult word”, rejected outright by Bjorvand & Lindeman (2000: 341-342)<sup>8</sup> and not even mentioned as a possibility in Kroonen (2013: 204). Besides, even if the connection should be correct, there are no obvious candidates for cognates outside the Northern European branches. No alternative etymologies are generally accepted. Orel assumes that it is in some way related to Skt. kálpate ‘succeed; fit; be partaken by’, kálpa ‘crossbeam’. This is rejected by Kroonen as semantically unconvincing; he simply states that the word has no clear etymology. Bjorvand & Lindeman likewise give up on most old proposals, vouching that the only possible solution is to operate with a derivative of the PIE root \*skel- ‘to split, share’ (probably \*skelH-) without s-mobile, an idea going back to Uhlenbeck (1905: 287). A serious issue is then that the root is unattested with s mobile – all attested forms have \*s- in the root.

<sup>7</sup> In Hyllested (2009: 204) I adhered to this etymology which I now reject. It goes back to at least Brugmann Grdr. (II, 1: 388-389) although indirectly ascribed to von Grienberger (1900: 107-108) by Orel (2003: 154) who gives a summary of the classical proposals.

<sup>8</sup> I do not quite understand that their rejection seems to be based on the difference in semantics; cf. that the derived substantive \*halbōn- f. means ‘side’ (Ofris. halve, OS, OHG halba) and ‘region, quarter’ (ON hálfr), to a great extent overlapping with the Lithuanian semantics.

It is important to take into account the semantic scope of the word which can be inferred from its use in the attested languages. PGmc. *\*halba-* seems to have three (related) meanings: '1. incomplete, partial, reduced'; 2. 'a significant portion of'; 3. 'exactly ½'. There is no particular reason to think that the latter meaning, historically speaking, is the (only) primary one – numerals and terms related to measurement most often originate from words with broader meanings.

For Proto-Balto-Fennic a word *\*halpa-*, gen. *\*halþan-* can be reconstructed with the meaning 'reduced, cheap of prices'. The reflexes in the individual Balto-Fennic languages are Fi. *halpa*, gen. *halvan*; Est. *halb*, gen. *halva* 'bad, evil; poor (e.g. of prospects, arrangement); ill; low (of spirits)', SW and SE dial. also 'cheap (favorable of prices)'<sup>9</sup>; Votic *alpa* 'bad' (Adler & Leppik 1990: I, 99; attested only in folk tales). From Northern Balto-Fennic it has been borrowed into several of the Saami languages, e.g. N Saami *hálbi* 'cheap of price and goods', Lule Saami (h)*alpō-* 'become cheaper (of goods)'.

Koivulehto (1982; via Hofstra 1985: 162) has proposed that it is a borrowing from Proto-Germanic into Fenno-Saami (= Early Proto-Fennic), namely from the word *\*salwa-* 'filthy, dirty' (> ON *splr* 'dark grey') which would be manifested in Fenno-Saami as *\*šalwa-*, later yielding Late Proto-Fennic (= Balto-Fennic) *\*halpa-* ~ *\*halva-* where *\*halpa* is one example of *-h-* corresponding to PGmc. *\*-s-* in the oldest loans. Sammallahti (1998: 247-248) adds this etymology with a question mark.

However, Balto-Fennic *\*halba-* cannot be a loan from Germanic as it is cognate with Mari (Cheremis) Nw W *šul-δə*, E Ki U *šul-δo* 'cheap' (cf. UEW 782)<sup>10</sup> and thus come from a Fenno-Volgaic *\*šalV-* with the adject-

<sup>9</sup> Cf. further Est. *halvustama* 'to hold cheap, belittle'. Salaca Livonian *alu* 'cheap, favorable (of prices)' has been borrowed from from SW Estonian *alv* (Kettunen 1938: 9).

<sup>10</sup> Proto-Mari *\*-u-* is not the most common outcome of Proto-Fenno-Volgaic (or Proto-Uralic) *\*-a-*, which is normally *-a-* ~ *-o-* depending on the dialect. UEW (782) mentions, however, that there are more examples of Mari *\*-u-* < Fenno-Volgaic *\*a-*. We can at least find examples of Balto-Fennic *\*-a-* corresponding to Mari *\*-u-*, one of which is mentioned by Bereczki (1988: 338), namely Fi. *tammi* 'oak' ~ Mari O *tumo* W *tum*; but the question is which language group is conservative in this respect since Balto-Fennic seems to have many secondary *a*'s which might be due to a yet undetected conditioning. In other words, the correct Fenno-Volgaic reconstruction might in fact be *\*šulV*. It is also conceivable that the Mari word was contaminated with *šulδa-(la-)* 'to cut a little bit' (cf. *šulδa-las* 'one cut across, one single cut'), a derivative of *šua-*, *šula-* 'to cut' which, however, itself may derive from a FU root with *-a-*, *\*šale-* 'to cut, split' (pace UEW 459-460). Most important for our purposes here is that the Balto-

tival and participial suffix *-pa* added to it. Accepting Koivulehto's etymology implies a rejection of the Mari cognates and the reconstruction of a Fenno-Volgaic term.

I therefore propose instead that the word was borrowed from Proto-Balto-Fennic into Proto-Germanic. Note that *\*halpa-*, even including its derived verb (Fi. *halventa* 'make cheaper') is already known to have been borrowed into Saami at a later stage. It is also important that there are no plausible cognates on the Indo-European side. Even if Mari cognates for *\*halpa* are not accepted, this would not hinder a Finnic → Germanic etymology.

Germanic must then have borrowed the oblique stem with weak grade (*\*-β-* before a closed syllable in the paradigm, gen. *\*halβan*) which explains the Germanic voiced fricative; the reason cannot be Verner's Law since the word would have had initial stress all along<sup>11</sup>. The Fennic word would have been borrowed first with the primary meaning '1. incomplete, partial, reduced', while the originally secondary meanings 2. 'a significant portion of'; 3. 'exactly ½' (which of course can overlap completely in practice, but do not have to) would be semantic specializations within Germanic, eventually replacing the original PIE word for 'half', *\*sēmi-*, *\*sm̥-tero-*. The motivation for the borrowing is clear: 'reduced' of prizes and 'cheeper' of goods would surely have been an important concept in trading contexts.

#### 1.4 PGMC. *\*PUNGA-* 'PURSE' ← BF *\*PUNKA-* 'CHEAP, REDUCED'

Turning to words with a different initial consonant, another mystery in Germanic etymology is *\*punga-* 'purse, pouch, small bag' (> Goth. *puggs* /pungs/, ON *pungr*, OE *pung*, ODu. *pung*, *pong*, MLG *punge* 'id.' [> Scandinavian], OHG *scaz-pfung* 'money-bag'). With its initial *\*p-*, it looks more than anything like a loanword. Most proposals seem to have just postponed the problem by suggesting as a source languages where it likewise has the air of a foreign origin (cf. Orel 2003: 293) – In fact, MLat. *punga*, Rumanian *pungă* and MGk. *ποῦγγα* are most likely to be Gothic loanwords (Bjorvand & Lindeman 2000: 707), and then we are back to square one. Neither they nor Orel (2003: 293) or Kroonen (2013)

---

Fennic and Mari forms can belong together and be traced back to a common form which is older than Proto-Germanic.

<sup>11</sup> Thus, Gmc. *\*-b-* next to *\*h-* here cannot be used as evidence for the Germanic sound shift having preceded Verner's law.

offer any solutions except that the two latter surmise that there is some kind of relationship with *\*pukōn-*. Bjorvand & Lindeman laconically state that it may very well be an old *Wanderwort*, but obviously wanderwörter have origins in specific languages, too, although they are harder to trace by definition. There is no reason *a priori* why some of the European wanderwörter would not turn out to have an origin in Uralic – or even be pure loanwords, borrowed in a bilateral process between a Uralic language and the target language.

Balto-Fennic had a word *\*punka* which goes back to Proto-Uralic *\*puŋka* (or *\*poŋka*) ‘swollen or expanded object’ (UEW 404, SSA 427) and has cognates in Saami, Permian and Ugrian languages. It is reflected in Fi. *punka* Est. *pung* ‘something chubby or protruding, clod, bump, swelling; leaf bud; bag, purse’. While the latter meaning in Estonian may well be due to Middle Low German or even later Swedish influence (Mägiste 1982-1983 [2000] VII: 2230-2231), the semantic development from ‘something swollen’ via ‘bag’ to ‘purse’ is straightforward, cf. Welsh *balleg* ‘purse’ < PIE *\*bʰl-no-* from *\*bʰel-* ‘swell’. Besides, Est. also has a variant *pong*, corresponding to Fi.dial. *ponki*, *ponka*, *ponko*, a variation in line with the one found in rest of Uralic.

I suggested already in Hyllested (2008: 136) that PGmc. *\*punga-* is a borrowing from Balto-Fennic, and, more tentatively, that the forms that look as if they go back *\*puh/kkōn-* (ON *poki* ‘bundle, purse’, OE *pocca*, *pohha* ‘id.’, Orkney Norn *buggy* ‘belly’, MDu. *poke* ‘bag for wool’, NHG *Pfoch* ‘bag’) may in fact be borrowings from a Western Saami precursor of Mod. N Saami *buggi* ‘bump, lump, hump; swollen or expanded object’ and/or *boggi* ‘short fat one (person, animal or thing)’. The medieval assimilation of nasal plus stop (*\*-NT-* to *\*-TT-*) in Western Saami serves as *terminus post quem* since this development took place after 1000 AD and spread to Central Saami not until the late 1500’s (Sammallahti 1998: 29, Hyllested 2008: 134). The transmission can have taken place in several steps, most likely with Nordic as the middleman. Note that the words *bag* (< ME *bagge*) and *pack* itself is already a borrowing from Western Saami *\*pakke-* (> N Saami *baggi*) via Old Norse *baggi* (Hyllested 2008: 136) ‘package, bundle; plump animal’. Medieval Saami and Balto-Fennic material in Norse and West Germanic relate typically to hunting and fishing, the fur trade, the production of hide and down, and trade concepts in general. Terms for ‘purse’ or ‘bag (made of hide)’ – with emphasis on either meaning – obviously fit in this picture.

## References

- Adler, Edna & Merle Leppik, 1990: *Vadja keele sõnaraamat*. Tallinn: Keele ja Kirjanduse Instituut.
- Bereczki, Gábor, 1988: "Geschichte der Wolgafinnischen Sprachen". – Denis Sinor (ed.): *The Uralic Languages. Description, History, and Foreign Influences*. Leiden / New York / Copenhagen / Cologne: E.J. Brill. Pp. 314-350.
- Bjorvand, Harald & Fredrik Otto Lindeman, 2000: *Våre arveord*. Oslo: Novus.
- von Grienberger, Theodor, 1900: *Untersuchungen zur gotischen Wortkunde*. Vienna: Gerold.
- Hofstra, Tette, 1985: *Ostseefinnisch und Germanisch. Frühe Lehnbeziehungen im nördlichen Ostseeraum im Lichte der Forschung seit 1961*. Groningen: Rijksuniversiteit Groningen.
- Hyllested, Adam, 2008: "Saami Loanwords in Old Norse". – Hans Frede Nielsen (ed.), *Early and Prehistoric Language Development in North-Western Europe* [= NOWELE 55/56]. Odense: University of Southern Denmark Press. Pp. 131-146.
- Hyllested, Adam, 2009: "PIE \*-bʰ- in Nouns and Verbs: Distribution, Function, Origin". – Rosemarie Lühr & Sabine Ziegler (eds.): *Protolanguage and Prehistory. Akten der XII. Fachtagung der Indogermanischen Gesellschaft, vom 11. bis 15. Oktober 2004 in Krakau*. Wiesbaden: Reichert. Pp. 202-214.
- Kettunen 1938: *Livisches Wörterbuch* [= V]. Helsinki: Suomalais-Ugrilainen Seura.
- Koivulehto, Jorma, 1982: "Zur Periodisierung germanischer Lehnwörter im Ostseefinnischen; wie findet man germanische Lehnwörter und was sagen sie uns? Lecture given at Rijksuniversiteit Groningen 12 and 13 May, 1982.
- Kroonen, Guus, 2013: *An Etymological Dictionary of Proto-Germanic*. Leiden: Brill.
- Mägiste, Julius, 1982-1983 [2000]: *Estnisches etymologisches Wörterbuch*. Helsinki: Suomalais-Ugrilainen Seura. 2<sup>nd</sup> ed. 2000.
- Sammallahti, Pekka, 1998: *The Saami Languages: An Introduction*. Karasjok: Davvi Girji.
- SSA = Ulla-Maija Kulonen & al. (eds.): *Suomen Sanojen Alkuperä. Etymologinen Sanakirja*. I-III. Helsinki: Suomalaisen Kirjallisuuden Seura and Kotimaisten Kielten Tutkimuskeskus. 1995.
- Uhlenbeck, Christianus Cornelis, 1905: "Bemerkungen zum gotischen Wortschatz". – *Beiträge zur Geschichte der deutschen Sprache und Literatur* 30: 252-327.
- Zoëga, Geir T., 1910: *A Concise dictionary of Old Icelandic*. Oxford: Clarendon.





## Gothic *mammo* ‘meat’ in the Light of Saami Evidence

### Abstract

The Gothic hapax *\*mammo* [f.] ‘meat’ reflects Pre-PGmc. *\*mamz-mōn* ‘flesh’, which goes back to PIE *\*moms-mo-*, a body-part derivative of *\*mems-* ‘meat’. Decisive evidence for this exact formation is delivered by North Saami *muošmi* ‘meat between the thigh and ribs (of a reindeer)’ which is not a Palaeo-Laplandic substratum word, but a regular substitution of the Pre-Proto-Germanic word.

### 1 Gothic *mammo* – attestation and earlier proposals

Gothic *\*mammo* [f.] ‘meat’ is a hapax legomenon occurring in the gen.sg. *mammons* from the Epistle of Paul to the Colossians, 1, 22<sup>1</sup>:

*in leika mammons is pairh dauþu du atsatjan izwis weihans jah un-*  
*wammans jah usfairinans faura imma*

νυνὶ δὲ ἀποκατήλλαξεν ἐν τῷ σώματι τῆς σαρκὸς αὐτοῦ διὰ τοῦ  
θανάτου παραστήσαι ὑμᾶς ἁγίους καὶ ἀμώμους καὶ ἀνεγκλήτους  
κατενώπιον αὐτοῦ

“In the body of his flesh through death, to present you holy and un-  
blameable and unproveable in his sight”

---

<sup>1</sup> The Gothic translation of the relevant passage is handed down to us via Codex Ambrosianus A and Codex Ambrosianus B, both of which are now kept in Biblioteca Ambrosiana in Milan.

Uhlenbeck (1905) originally suggested that the word had arisen as pure nursery language via the meaning ‘mother’s breast’ like Lat. *mamma* ‘breast’, but a year later (cf. Uhlenbeck 1906) he had abandoned that viewpoint. Nonetheless it was maintained by Scardigli many decades later (1973: 72).

Today most scholars regard *mammo* as an o-grade formation based on PIE *\*mēms-ó-* (? *\*mems-ó-*) > PGmc. *\*mimza-* [n.] > Goth. *mimz* ‘meat’. This was first suggested by Mikkola (1897) who at the same time left the possibility open that it could be derived from a PGmc. *\*mazmōn* based on the root in OHG *muos* ‘food’. His etymology was admittedly received with some hesitation by the standard etymological dictionaries of Gothic, e.g. Lehmann (1984: 243) who earnestly mention the possibility without directly endorsing it. Lehmann tags the label “disputed” on it, and Casaretto (2004) loyally maintains that the etymology is “unclear”.

Things did not start to develop until Polomé (1967). Kroonen (2013) express similar thoughts in his new dictionary, stating that *mammo* can reflect either

- a) Pre-PGmc. *mamzōn-* < *\*moms-ó-* with an uncertain assimilation of *\*-mz-* > *\*-mm-*, or
- b) Pre-PGmc. *\*mamz-mōn* < *\*moms-mó-* (*\*moms-mén-*) involving the well-known assimilatory development *\*-zm-* > *\*-mm-*.

He notes that the same problem is relevant for PGmc. *\*kramma-* ‘moist, humidity’ (which correspondingly can come from either *\*kramzōn-* or *\*kramzmōn-*), embracing Polomé’s viewpoint that *\*-mzm-* is the more probable reconstruction on phonological criteria because it is uncertain whether an assimilation of *\*-mz-* into *\*-mm-* took place at all; the most obvious counterexample is PGmc. *\*mimza-* itself<sup>2</sup>.

<sup>2</sup> On this basis, the reconstruction has been characterized as problematic by Grienberger (1900: 154), Trautmann (1906: 62), Falk/Torp (1909: 310), and Casaretto (2004: 235), but although the sequence must be morphologically segmented *\*-mz-m-*, it of course forms a phonological string *\*-mzm-*, and while it is true that *\*-z-* generally is not assimilated to a preceding *\*-m-* in Gothic or Germanic in general, it does assimilate to a following *\*-m-*. Thus, reconstructing *mammo* as *\*mamz-ma-* only appears problematic if the formation is supposed to have taken place in Gothic itself (when the assimilation process was no longer operating); but there is no need to think that it did not happen in Germanic or even PIE already.

From a morphological perspective, however, *\*mamz-mōn* is more difficult to understand: It is the only attestation of *\*mems-* ~ *\*mēms-* with an o-grade anywhere in Indo-European. And one wonders what the motivation would have been for forming a new *\*-mo-/\*-men-* stem from a nominal (pseudo-)root which already contains two labial nasals.

## 2 North Saami *muošmi* and its implications

Turning to neighboring languages, new evidence has appeared. The North Saami word *muošmi* means ‘meat between the thigh and ribs (on a reindeer)’<sup>3</sup> and has no generally accepted etymology. Aikio (2012) on the basis of its phonological structure deems it as a loanword. Moreover, he includes it in a long list of reindeer terminology inherited from the original language of the present-day Saami populations, called “Palaeo-Laplandic” (he explicitly avoids *Saami* here in order to distinguish it from modern Saami language and culture). It is generally acknowledged that the Saami populations were not speakers of Uralic from the outset but at some point took over a Uralic language, and that an abundance of terms from the old sunstratum language survived, accounting for hundreds of words which cannot be inherited nor borrowed from any known language group. Aikio is the forerunner in the identification and analysis of this substratum; another important work in this field is his essay from 2002 where *muošmi* however is not mentioned.

It has been known for much longer that Saami languages, like Balto-Fennic, have borrowed hundreds of words from Germanic through different periods in history, and that there are also Pre-Proto-Germanic loanwords, i.e. items whose present-day Saami shapes reveal that they were borrowed before the sound changes that define Proto-Germanic had all taken place. While it is beyond doubt that a large part of the Saami lexical stock can be ascribed to one or more non-Uralic and non-Indo-European substrate languages, we should therefore not be surprised if at least a small part of the proposed substratum words turn out to be Germanic or Pre-Germanic after all.

The Proto-Saami form of *muošmi* would mechanically be reconstructed as *\*masmV-*. However, *\*mamsmV-* is in fact also possible: The normal development of Proto-Saami *\*-amC-* is undoubtedly North

<sup>3</sup> Nielsen & Nesheim (1932-1965, bd. II 1934 by Nielsen only: 702) gives the more elaborate definition: ‘layer of fat and muscles between the thigh and ribs of reindeer and other animals (goes with the side when the carcass is cut up)’.

Saami *-uovC-*, cf. *ruovda* ‘border, edge’ ← Pre-PGmc. *\*ramþa-* (*\*ramða-*) > PGmc. *\*ranþa-*. However, the sequence *-uov-* is apparently phonotactically illicit before *\*-š-* (with or without a consonant following), meaning there are simply no words in contemporary North Saami containing the structure †*-uovš-*. I could find none in the extensive Nielsen & Nesheim (1932-1962).

Although *\*muovšm-* could very well have existed at some stage, Proto-Saami *\*mamsmV-* would regularly yield the North Saami end result *muošmi* which is also what we find. It therefore simply seems to be a Pre-Proto-Germanic loan from the very same word that yielded Gothic *mammo*. Semantically there are already no problems, but even the specific Saami meaning may reveal something about the history of the formation.

### 3 Motivation for formation in *\*-mo-*

One might argue that even if Pre-PGmc. *\*mamz-mōn-* or PIE *\*moms-mo-* was formed on the basis of *\*memzan-* or *\*mems-* respectively, the motivation for using exactly the *\*-mo-* or *\*-men-* suffix could still be based on the iconic nature of reduplicative semantics. Danish nursery terms like *mam(-mam)* and *mamse* ‘food’, although they clearly look as if they are formed like nursery words in any other language, are perceived by speakers as derived from the adult term *mad*, which may even partly be historically true (or impossible to determine). In other words, the use of specific morphological elements in word-formation, although formally allowed, can still be motivated by conditions outside the morphological system – such as association to children’s language or even just a general human delight of playing with sounds and fascination of phenomena such as alliteration<sup>4</sup>.

<sup>4</sup> One might also speculate that Wulfila associated *\*mammo* with *\*mammona* ‘mammon, wealth’ (a masc. n-stem, only attested in the dat.sg. *mammonin*) in a religious context where the concept of ‘meat’ could be contrasted with spiritual properties and values (cf. “the hunger for wealth”). However, it is hard to imagine such an association among the heathen speakers of Pre-Proto-Germanic who would never have heard of any such Greek-Aramaic word, let alone the biblical concept it denoted. My guess is that if Wulfila felt an association between the two, the similarity of *mammōn-* f. to *mammōnan-* m. was simply a convenient coincidence of the kind that translators sometimes stumble upon and are happy to exploit for literal and pedagogical purposes (to the extent that they have a choice at all).

Still, in this case there is in fact a possibility that we might be dealing with word-formation based on a very clear semantical motivation. Quite a few Indo-European names for body parts occur with a suffix *\*-mo-* or *\*-men-*; although a few of them can be shown to have originally meant something else, e.g. substantivizations of adjectives (PCelt. *\*skama-* ‘lung’, but also the adj. ‘light’), they cannot all be explained in this way. Even when the term looks like productive deverbal nominal formation (such as PGmc. *\*barma-* ‘bosom’ < PIE *\*b<sup>h</sup>er-* ‘to carry’) one may wonder why exactly *\*-mo-* was used in these cases. I think that there are enough examples for us to at least claim that its use was still in (marginal) use in PIE, and that it reflects an earlier grammatical marking of inalienable possession, using the 1<sup>st</sup> person verbal and pronominal marker *\*-m-*:

- PIE *\*g<sup>h</sup>eh<sub>2</sub>u-mo-* ‘gum’ (or *\*gau-mo-*; deverb. from *\*? \*g<sup>h</sup>eh<sub>2</sub>-* ‘gape’)  
 PIE *\*pleu-mōn* ‘lung’, Skt. *klóman-*  
 PIr. *\*čāš-man-* ‘eye’  
 Gk. δέρμα, Ru. *dermó* ‘skin’  
 Gk. ὄμμα ‘eye’  
 Gk. ὀφθαλμός ‘eye’  
 Gk. κνήμη, OIr *cnáim* ‘lower leg’, PCelt. *\*knā-mi-* ‘bone’  
 Lat. *abdōmen* ‘belly, abdomen’  
 Lat. *rūmen* (~ Skt. *romantha-* ‘chewing the cud’)  
 PCelt. *\*skama-* ‘lung’ (< adj. *\*skama-* ‘light’)  
 Alb. *zemër* ‘heart’ (< *\*-men*; root?)  
 PGmc. *\*armaz* ‘arm’ ~ Lat. *armus* ‘upper arm’ (< PIE deverb. from *\*ar-* ‘link’)  
 PGmc. *\*barma-* ‘bosom’ (< PIE deverb. From *\*b<sup>h</sup>er-* ‘to carry’)  
 PGmc. *\*parma-* ‘gut, intestine’ (< PIE deverb. ‘drill (?)’)  
 PGmc. *\*fapmaz* ‘bosom, lap’ (OHG *fadum*, OE *fæðm*; ON *faðmr* ‘spread arms, embrace; thread’) ~ PCelt. *\*patimā-*, *\*patamī-* ‘yarn, thread’ (> Gael. *aitheamh*, W *edafedd*), cf. Hamp 2008, Hyllested 2010  
 PGmc. *\*skama-* ‘shame; nakedness’ < ‘genitals, naked skin, flain skin, hairless skin’ (< *\*skā<sub>2</sub>-mo-* from *\*sekh<sub>2</sub>-* ‘cut’, cf. *sex*, Latin *secāre*)  
 PGmc. *\*skerma-*, *\*skermi-* ‘skin, hide’ (< *\*sker-* ‘cut (off)’  
 ON *hvarmr* ‘eye lid’, Far., Icel. *hvarmur*, Nw. *Kvarm*, Elfd. *ogen-walm* via Swedish < PGmc. *\*hwal(b)-maz*  
 German *Kieme* ‘gill’ (? ~ OHG *kiuwa*, *kewa*, *kouwa*)

This situation is parallel to Uralic where I believe to have uncovered a specific use of the denominal suffix for specifically for body-parts. The suffix seems to have stayed productive in several branches – at least it

must have lived longer than in Indo-European, judging by the impressive number of examples:

**Proto-Uralic (attested in both Fenno-Ugric and Samoyed):**

\*šilmä 'eye'

\*šedmä 'kidney'

\*šüδä-mV 'heart'

\*piŋsV-mV, \*pi(p)se-mV 'lip'

\*umV 'waist'

Proto-Saami \*kâjma 'marrow, brain' ~ Samoyed: Selkup *küümä*, Nenets *xæwa* 'marrow', Nganasan *kojmu* id. (~ Selkup *küün* 'marrow, brain')

**Proto-Fenno-Ugric (attested in both Fenno-Permian and Ugric)**

\*ñaŋkće-mV 'tongue, palate' (PU *ñaŋkće* 'tongue, palate')

\*ñälmä 'tongue'

\*čVjmV 'loin'

\*wiδV-mV 'marrow; brain'

\*ňarma 'groin' (formed from \*ňarV 'hairless skin', FU)

\*kel'mä, \*keδ'mä 'skin' (formed from \*keδ'e 'skin')

\*kul'ma 'corner of the eye; eyelid; temple'

**Proto-Fenno-Permian (attested in both Fenno-Volgaic and Permian)**

\*kurmV 'fist; handful'

**Proto-Fenno-Volgaic (attested in both Fenno-Saami and Mari-Mordvin)**

\*čilV-mV 'canine tooth'

\*sorme 'finger'

\*wajmV 'heart'

\*pijra 'crop (of animals)' has \*-mV in Saami and Mordvin

**Individual branches / languages**

PU \*piŋV 'fist' has \*-mV in Koibal Samoyed *pam*, Mator *ojme*

PU \*täwe 'lung' has \*-mV-ktV in Selkup *tyymäktä*, *tyymiektö*

PU \*künče 'nail' has \*-mV in Mator Samoyed *kadam*

PU \*konV + 2<sup>nd</sup> element 'arm pit' has \*-mV in Ostyak *xonəm pat*

PU \*täktä 'bone(s)' has \*-mV in Hung. *tetem*

PU \*kunV 'abdomen' has \*-mV in Komi *kynöm*

PU \*polwe 'knee' has \*-mV in Mordvin *pulma-*, *puma-*

PU \*kVŋkV 'Adam's apple' has \*-mV- + -s- in N Saami *guoggom(as)*

PU \*ajŋe (\*ajwe) 'brain' has \*-mV in Lule Saami *vuojńam* < PS \**vuojŋēm*

PPF \*čVIV 'small finger' has \*-mV in West Saami \*čVIV-čV-mV (the base with \*-čV- is common Saami)

PU \*ikene 'gums, gills' has -ne replaced by -me in Fi.dial., *ikemet*, Kar. *igimet*, Vot *ičemet*, Est. *igemed*, *igem*, *igim*, *ikem*, *ikim*, Liv. *i'gmōd* (in Est. and Liv., only forms with -m- occur; nom. -n can come from both consonants)

Estonian *habe(me)* ‘beard’  
 Finnish *siera(-im-(inen))* ‘nostril’  
 Finnish *helma* ‘lap, bosom’  
 Votic *süamiüs* ‘inner parts of animal’s body’

#### 4 Conclusion

Although Goth. *mammo* is normally translated as ‘meat’, our knowledge of it is based on a single attestation where it clearly refers to the human flesh. And N Saami means not ‘meat’ as a kind of prepared food, but exactly a kind of ‘flesh’, i.e. meat as a certain part of the animal’s body.

Not all languages have the same distinction as English, but we may conveniently use English to illustrate a word-formation scenario where *\*-mo-* transformed the meaning from ‘meat’ in general into ‘flesh’ specifically:

PIE *\*mems-* ‘meat’  
 + body-part term suffix *\*-mo-*  
 →  
 PIE *\*moms-mo-* (with body-part term suff. *\*-mo-*) ‘flesh’<sup>5</sup>

Since the suffix does not appear to have survived in any of the Indo-European subgroups, I prefer to assume that the formation took place in PIE itself. However, most examples of Indo-European body-part terms with *\*-mo-* are Germanic so we cannot exclude that it was still productive at least at the Pre-Proto-Germanic stage where *\*mamz-mōn-* would have been formed (perhaps first as *\*mamz-mō*). This word was borrowed into Proto-Saami as *\*mamsmV* and developed ultimately into North Saami *muošmi* (perhaps via *\*muovšmī*) but disappeared from the other Saami subgroups.

It perhaps appears surprising that a word belonging to reindeer terminology was borrowed from (Pre-)Proto-Germanic into Proto-Saami. Judging by the number of reindeer terms that survived from Palaeo-Laplandic into Saami, one might at first glance conclude that reindeer breeding was important already in Palaeo-Laplandic times. Domestica-

<sup>5</sup> The transition of a simple thematic feminine *\*-eh<sub>2</sub>* into the productive weak class of feminines in *\*-ōn* in Germanic is of course well known, but has not yet been assigned a satisfactory explanation. Schmidt (1985) mentions the possibility that the gen.pl. *\*-ōn* having arisen could have provided a motivation for analogy; see also Thöny 2013: 119.



tion of reindeer may have taken place in the transition period between the North European Bronze Age and the Iron Age (1200-600 BC). However, it is worth noticing that most Palaeo-Laplandic terms do not specifically refer to breeding; it could just as well reflect a way of life characterized by reindeer hunting.

### References

- Aikio, Ante, 2004: "An essay on substrate studies and the origin of Saami". – Irma Hyvärinen, Petri Kallio & Jarmo Korhonen (eds.): *Etymologie, Entlehnungen und Entwicklungen: Festschrift für Jorma Koivulehto zum 70. Geburtstag* [= *Mémoires de la Société Néophilologique de Helsinki* LXIII]. Helsinki: Société Néophilologique. Pp. 5-34.
- Aikio 2012 = Luobbal Sámmol Sámmol Ante (Ante Aikio): "An essay on Saami ethnolinguistic prehistory". – Riho Grünthal & Petri Kallio (eds.) *A linguistic map of prehistoric Northern Europe* [= *Mémoires de la Société Finno-Ougrienne* 266]. Helsinki: Suomalais-Ugrilainen Seura 2012. Pp. 441-451.
- Casaretto, Antje, 2004: *Nominale Wortbildung der gotischen Sprache*. Heidelberg: Winter.
- Feist, Sigmund, 1939: *Vergleichende Wörterbuch der gotischen Sprache*. 3rd ed. Leiden: Brill.
- von Grienberger, Theodor, 1900: *Untersuchungen zur gotischen Wortkunde*. Vienna: Gerold.
- Hamp, Eric, 2008: "Germanic \**fapmaz* and Gravity in the North". – *Early and Pre-historic Language Development in North-Western Europe* (NOWELE 54/55), 349-52.
- Hyllested, Adam, 2010: "The Precursors of Celtic and Germanic". – Stephanie W. Jamison, H. Craig Melchert & Brent Vine (eds.) Review of Grünthal & Kallio. – *Kratylos*. Pp. 63-117.
- Kroonen, Guus, 2013: *An Etymological Dictionary of Proto-Germanic*. Leiden: Brill
- Lehmann, Winfred P., 1986: *A Gothic Etymological Dictionary. Based on the 3<sup>rd</sup> ed. Of Vergleichendes Wörterbuch der Gotischen Sprache bei Sigmund Feist*. Leiden: Brill.
- Mikkola, Joos J., 1897: "Baltische Etymologien". – *Beiträge zur Kunde der indogermanischen Sprachen* 22: 239-254.
- Nielsen, Konrad & Asbjørn Nesheim, 1932-1962 (bd. I-IV by Konrad Nielsen only): "Lapp dictionary based on the dialects of Polmak, Karasjok and Kautokeino / Lappisk ordbok grunnet på dialektene i Polmak, Karasjok og Kautokeino" [= Instituttet for Sammenlignende Kulturforskning, Serie B: Skrifter XVII]. Oslo: Aschehoug & co. (W. Nygaard). I-VI.
- Polomé, Edgar, 1967: "Notes on the Reflexes of IE /ms/ in Germanic". – *Revue Belge de philologie et histoire* 45, 3: 800-826
- Scardigli, Piergiuseppe, 1964: *Lingua e storia dei Goti*. Florence: Sansoni.

- Schmidt, Gernot, 1985: “Der Genetiv der indogermanischen *ā*-Deklination”. – Hermann M. Ölberg & Gernot Schmidt (eds.): *Sprachwissenschaftliche Forschungen. Festschrift für Johann Knobloch*. Innsbruck. Pp. 393–402.
- Thöny, Luzius, 2013: *Flexionsklassenübertritte*. Innsbruck: IBS.
- Trautmann, Reinhold, 1906: *Germanische Lautgesetze in ihrem Sprachgeschichtlichen Verhältnis*. Kirchhain.
- Uhlenbeck, Christianus Cornelis, 1905: “Bemerkungen zum gotischen Wortschatz”. – *Beiträge zur Geschichte der deutschen Sprache und Literatur* 30: 252–327.
- Uhlenbeck, Christianus Cornelis, 1906: “Aantekeningen bij gotische etymologieën”. – *Tijdschrift voor Nederlandse Taal- en Letterkunde* 25: 243–306.



## **The Mysterious Elder: Common Traits in European Names for *Sambucus nigra* and *Viburnum opulus***

### **Abstract**

No PIE names for the elder species (*Sambucus*) or the smaller water elder (*Viburnum opulus*) can be reconstructed although they are indigenous in all of the possible homeland areas. However, Lat. *sambucus* can be traced back to a meaning ‘the dusty tree’, referring to powdery mildew on canes, leaves and berries which gives the tree an overall dusty impression. This meaning also lies behind Lith. *šeivā-medis* (cf. *šývas* ‘grey; mildew’), a relic of the old Balto-Slavic word which was borrowed into Fenno-Permian as *\*šewa*. In common Slavic it was replaced by the loan *bъzъ*, *buzъ* < Old Turkic *boz*, *buz* ‘grey’ (cf. Ru. *búsel* ‘mildew’) while in the NW Slavic fringes the old name only still survived at least until the 20<sup>th</sup> c. in the amalgamated and folk-etymologically reshaped Sorbian form *dźiwi bóz* and as a loanword in neighboring German dialects (*Schibchen*, *Ziwecken*). Berries of the water elder are called *\*šarša* in Mari (Cheremis) which can be traced back to a loan from (East) Baltic *\*šeršas* ‘moldy’. Germanic words for ‘elder’ correspondingly can be analyzed as containing PIE *\*pelH<sub>1</sub>*- ‘grey’.

The frequent renamings of both species all over Europe may be due to taboo, linked to popular beliefs of the elder being guarded by “The Elder Mother” and her abilities to ward off evil. She was associated with spinning, considered a partly magical activity with links to the other-world, which triggered a folk-etymological reshaping of the Balto-Fennic forms after words for ‘thread’. The denotation ‘the grey one’ could have had a double connotation ‘elder tree’ and ‘old lady’, remarkably alike the homonymy covered by Mod. Eng. *elder*.

### 1 The elder is no beech: Lat. *sambucus*

As is well known, the Indo-European word for ‘beech’ has played a very important role in scholarly attempts to locate the Proto-Indo-European homeland, not least because the distribution of *Fagus sylvatica* at the period in question was known to have been restricted to areas to the West of a line from Kaliningrad to Odessa. For decades the classical reconstruction of the ‘beech’-word was PIE *\*b<sup>h</sup>āg-o-* ~ *\*b<sup>h</sup>auḡ-o-* (or *\*b<sup>h</sup>euḡ-o-*), although the mechanisms behind such alternations were not really understood – they were just thought to be reflected in the material. This outdated reconstruction can still be encountered once in a while. At the same time, the ‘beech’-word served as the eternal parade example of how the reconstruction of a homeland and names for species often rely on specific denotations which may have changed as populations migrated to new areas.

While most scholars nonetheless agreed on a PIE meaning ‘beech’ on the basis of Latin *fāgus*, PGmc. *\*bōkō-*, *\*bōkjōn-*, and Continental Celtic *\*bāk-* in Gallo-Lat. *Bacena silva* ‘beech forest’, they also had to consider a) Greek φηγός ‘Kermes oak’ which phonetically corresponds to the Western European forms, but has acquired a new meaning; b) Albanian *bungë* which agrees with Greek semantically but contains a mysterious nasal; c) Kurdish *buz* ‘elm’ and d) the Slavic word for ‘elder’, *\*bъzŭ* (Cz. *bez*, Bulg. *băz* etc.)<sup>2</sup> or *\*buzŭ* ‘elder’ which combined several problems of phonology – reflex of both *u*-vocalism not present elsewhere and a palatal stop not matching the Albanian word – as well as the fact that they designated other trees (to the extent that the elder can be called a tree at all). Furthermore, Germanic with its reflex of an unaspirated voiced stop seemed to be incompatible with the lack of Winter’s law in Slavic<sup>3</sup>.

Scholars should have put more weight on the suspiciously sparse occurrence in Eastern languages in general and Satem languages in particular – the term seemed to be absent in Baltic, Armenian, Indic, and most Iranian, as well as Anatolian and Tocharian). An excellent histori-

<sup>1</sup> Probably just a regular development of *\*b<sup>h</sup>ag-n-* (cf. my article on Alb. *hundë* elsewhere in this publication).

<sup>2</sup> Blažek 2002: 201-202 mentions the need to operate with an analogous *\*baz* on the basis of the genitive *\*baza* to account for Bulg. *bās*, SCr. *baz*. However, this is unnecessary since these forms come regularly from *\*bъzъ*.

<sup>3</sup> Kroonen (2013) mentions that *beyki* can just be an irregular continuant of ON *bæki*, a form directly continued by Icel. *bæki*.

cal overview as well as detailed and important considerations based on lots of new evidence, including from satem languages and Anatolian, can be found in Blažek (2002 with references). His most important conclusion for purposes here is that, based on the last 50 years of research, no evidence remains for any irregular alternations in the word. In fact, before Blažek himself, Mod. Icel. *beyki* is the only form reflecting *-u-* that is associated with the beech. He argues that the Icelandic word is not inherited at all, but a 17<sup>th</sup> century rendering of Dan. *bøg* with pseudo-etymological restitution of *-ey-* for Da. *-ø-* (as in inherited words); he adds the not important factor that no beech forests, or even any kind of proper forests, are found in Iceland<sup>4</sup>. Henning (1963): *buz* ‘elm’ is not related at all, but instead a regular Central Kurdish reflex of PIE *\*ueiǵʰ-* (corresponding to *wiz* in other Kurdish dialects; Slavic *\*vežĩ*, Eng. *witch*). Blažek shows that there Iranian reflexes showing exactly *-āg-* namely Gilaki *fay*.

Slavic *\*bъzъ*, *\*buzъ* (also *\*buzina*, *\*bъzina*, cf. Ru. *buzinā*) is probably not related. Quoting Blažek, “it is surprising how many excellent scholars were ready to accept the relation between *\*bʰāg(o)-* and the Slavic forms indicating *\*b(h)euǵʰ(h)-* (...) in spite of the difference in root vowel and in semantics”. One might add that, even even for the excellent scholars mentioned, the Slavic ‘elder’-word remained the only clear indication of a) an irregular variation between *-ā-* and *-u-* and b) the palatal character of the following stop. Based on typological evidence, Blažek himself suggests that *\*bъzъ*, *\*buzъ* is derived from PIE *\*bʰuǵʰ-* ‘he-goat’ (Eng. *buck*; this is the case e.g. in Caucasian languages). He considers Lat. *sa(m)būcus* ‘elder’ a loan from Gaulish, formed with an alleged IE tree-name prefix also found in e.g. Gk. *σμῖλαξ* ‘yew’ ~ Lat. *īlex* ‘holm oak, kermes oak’.

While I agree completely with Blažek that the Slavic word should not be included in the ‘beech’ word-family, I have difficulties accepting his alternative etymological proposal, mainly because of the structure of the Latin word. It cannot be ignored in this context that *sa(m)būcus* clearly appears to contain the well-known plant-name suffix *-ūcus*, *-ūca* as in *albūcus* ‘asphodil’, *lactūca* ‘lettuce’, deriving from PIE *\*-u-h<sub>3</sub>ʷo-* ‘looking like X’<sup>5</sup>. Moreover, the Latin words displays a variation between

<sup>4</sup> Kroonen (2013) mentions that *beyki* can just be an irregular continuant of ON *bæki*, a form directly continued by Icel. *bæki*.

<sup>5</sup> Thus unrelated to *sambūcus* ‘harp, flute’ which, like the synonyms *ambūbāia*, Gk. *ἰαμβύκη*, *σαμβύκη* is a later loanword from popular Aramaeic with secondary *-mb-* from *-bb-*, cf. Literary Aramaeic *sabbākā*, Syr. *‘abbubā* (see

*samb-* and *sab-* as mentioned by Blažek himself: The form *sabūcus* is also attested. Unless we are dealing with a folk-etymology, it therefore seems that we are dealing with an inner-Latin formation to a stem *samb-* alternating with *sab-*. We know one word which shows such alternation, and that is *sabulum* 'sand', although forms with a nasal are not otherwise found within Latin. Before we consider possible links between the elder and the meaning '(looking like) sand', let us return to the Slavic word.

### 3 The Turkic origin of Slavic \**buzь*, \**bъzъ*, Mod.Gr. *μπούζα*

In my opinion the best candidate for the Slavic word is a borrowing from an early Turkic language. OTurk. *boz* 'grey' and its relatives are already known to be the source of numerous words in East Slavic, not least Old Russian, displaying various manifestations of the original vowel and sibilant – and notably a word among them meaning 'mildew'. One of them also occurs in West Slavic (examples from Vasmer, Räsänen 1969, Dyneley Prince 1919).

- ORu. *busyj*, *bosyj* 'grey', *busétъ*, *busovétъ* 'to become grey, blue or dark', cf. O.Ru. *bosym volkom* 'as a grey wolf' (in the Tale of Igor's cam paign) and *busovi vrani* 'grey crows'  
 Mod.Ru. *buzlák* 'Crocus reticulatus'  
 Mod.Ru. *buzlúk* 'traces of ice on the boots' (Tat. *bozluk*; *boz* 'ice')  
 Mod.Ru. *buzán*, *búsol*, Ukr. *busel*, Belaru. *buseł* 'white stork'  
 Mod.Ru. *buzún* 'salt from salt lakes'  
 Mod.Ru., Ukr. *búsel* 'mildew'  
 Pol. *busieł*, *buško*, *buś*, *busek* 'young stork'

The reason for the large representation of this Turkic word as a loan-word in Russian is no doubt the importance it has played in the history of Turkic and Mongolian identity from the times of the ancient steppe cultures to the present day, mainly connected with its function as epithet of the wolf which functioned as a totem animal for the early Turkic peoples, cf *boz kurd*, *boz börü*. Correspondingly, in the Secret History of

---

Schwyzler 238-239 with numerous parallels). Brück's attempt to relate it to Lat. *scabō* must be rejected, although we cannot exclude his claim that *sabūca* could have been reshaped to *sambūca* under the influence of *sambūcus*.

the Mongols, the forefather of Temujin is a grey wolf called *Börte* Chino (born 758), with *börte* meaning ‘grey-blue, grey-white’.

These are all East Slavic, but Late Common Slavic bordered areas inhabited by Altaic tribes (Birnbaum 1998, Andersen 2003), and the etymon is also known from the Bulgarian grey drink *buzá* ‘a grey kvas-like drink’, borrowed by Turkish and perhaps the source of Eng. *booze* via Romani (cf. also Chagatai, Osmannic Turkic etc. *boza* ‘drink made of camel’s milk’ and Chuv. *pora*, its r-Turkic counterpart, which may ultimately be the source of the Gmc. *beer*-word). Quite remarkably, Mod.Gr. *μπούζα*, obviously a late loan, means ‘water elder’ – either it is a South Slavic loan although South Slavic forms seem to reflect only *\*bъzъ* – or it is directly from Turkish with a meaning not attested there, having replaced the name for the same plant as in Slavic earlier in history.

The final Slavic *-ъ* does not historically represent a vowel, but is just the automatic LCS (and OCS) manifestation of a word-final non-palatalized consonant in loanwords, cf. LCS *\*klobukъ* ‘hat’ (< Turkic, cf. Crimean Tatar *kalpak* ‘cap’), OCS *kovъčegъ* ‘box, casket’ < supposedly Avar; cf. Mongolian *qagurčag*), and LCS *\*tълmačъ* ‘interpreter’ (< Turkic *dolmač*).

#### 4 Lithuanian *šeivā-medīs*, West Slavic *\*šiv-*

Interestingly, a meaning ‘grey’ turns out to be compatible with a Baltic word for the elder. Lith. *šeivā-medīs* ‘elder’ a compound that can be analyzed synchronically as the ‘spool’-tree, because of its hollow branches that can be used as bobbins on a loom, cf. *šeivà* (cf. also *šeivā-kaulis* ‘radius (bone)’ next to *šeivī-kaulis* and *šeivā-filmė* ‘spool film’). However, one of the established PIE reconstructions for ‘grey’ is the formally identical adjective *\*keiH-uo-* ‘grey’ whose zero-grade *\*kiH-uo-* occurs in Lith. *šývas* ‘white; light grey; “moldy” (mostly of cows)’, *šývis*, *šývis* ‘mildew’, OPr. *sywas* ‘white’, Slavic *\*sivъ* ‘dark grey’. Lith. *šeivā-* in *šeivā-medīs* simply looks like the feminine counterpart *\*keiH-ueh<sub>2</sub>*. Gliwa (2008) derives the name of the tree exactly from *šeivà* ‘spool’ but he regards the etymology presented here as equally possible (p.c.).<sup>6</sup> One im-

<sup>6</sup> Admittedly, it must be noted that there are parallels to the development ‘spool’ > ‘elder’. At least Wichmann, Uotila & Korhonen (1987: 231) note that in the old handwritten Udmurt-Russian dictionary by Islenčev, it is noted that in the Jela-buga dialect, *šerj* has come to mean *Sambucus*. It is a borrowing from Chuvash



portant fact is that the name *šeivā-medīs* not only denotes the elder proper (*Sambucus* ff.) but also the so-called water elder, also called dwarf elder or danewort (*Viburnum opulus*) which has red berries.

The designation is probably the old Balto-Slavic one which can be inferred partly from a loanword in Permian languages denoting the water elder only (see below), partly from what look like relics from the Northwest Slavic fringes which notoriously retain archaisms from even before the Late Common Slavic period. In the German dialect of Upper Saxony the elder is called *Schib-chen*, and the Mansfeld dialect has *ziwe-cken*; both forms must belong with Sorbian *džiwi bóž*, which is synchronically ‘wild elder’ (Brüch 233, fn. 1), but probably a folk-etymology; it must be based on the use of the term *wilder Flieder* in NE German dialects that arose in order to be able to distinguish the locally preserved term *Flieder* ‘elder’ from that of the lilac, a new meaning that *Flieder* has acquired in High German after the introduction of the lilac into the country in the 16<sup>th</sup> c. (Kluge-Seebold 1989: 220). Sorbian dialectal convergence could easily have caused confusion between initial postalveolar sibilants and affricates even without the folk-etymology (Mucke 1891) but we cannot exclude that the similarity is a coincidence if the Sorbian loans in German dialects are misunderstandings as asserted by Brüch. With less certainty than for the Baltic and Permian attestations, then, we can reconstruct a West Slavic *\*šiv-* ‘elder’.<sup>7</sup>

For PGmc. we can reconstruct *\*flīpra-* ~ *flīðra-* on the basis of definitely a derivative with the three-name suffix *\*-ðra*. It is now feasible to assert that the first part is the other widespread PIE adjective meaning ‘grey’ (probably originally a different nuance) *\*pelh₂-*, known from e.g. Slavic *\*plěšni* ‘mold’, Lith. *pilkas* ‘grey’. The zero-grade was originally *\*pliH-* which should yield PGmc. *\*flī-*, but alternative zero-grades exist (Lit. *pilkas* < *\*pl-*) which might either indicate analogy from formations where the laryngeal had merged with the suffix (e.g. participial *\*pliH-to-* > *\*pli-tʰo-*) or that the laryngeal and perhaps even the *\*-i-* are extensions of an originally shorter root *\*pel-*, *\*plej-*.<sup>8</sup>

---

*šörö*, *šəra* ‘spool’ into Votyak from where it has been borrowed further into Komi as *šurj*; here, it reportedly only means ‘spool’.

<sup>7</sup> It is perhaps also worth noticing that the old o-grade possibly occurring in the Lith. dialectal form of ‘spole’, *šaiṽa*, and at least in Latv. *saiva*, does not occur in the tree-name (although that would be formally possible). The tree-name also occurs as *šeivmedīs*.

<sup>8</sup> Kroonen on the basis of a dialectal Dutch (N Hollandish) form *vlaar* reconstructs Old Frisian *\*fliar* and thereby PGmc. *\*fleupra-*, but considering how many Dutch forms are reshaped after verbs for ‘to bævre, flagre’ (cf. *vlinder*,

## 5 Why ‘the grey one’?

At first glance it seems odd that a name for several species of edible tree-like plants, which, depending on the season, are dominated either by an impression of shining black or red berries or bright white flowers would be designated by a color adjective ‘grey’. We can consider the possibility of a semantic shift from either:

- a) The white flowers of *Sambucus nigra* and other species
- b) The white pith of the elder, a traditional material for angli
- c) The black berries of *Sambucus nigra* (cf. NHG *Holunder* ‘elder’ and Ru. *lalína* ‘water elder’ < PIE *\*kel-* ‘black’ which already reveals that this species and not e.g. *Viburnum opulus* with its red berries was the first plant to bear this name), or
- d) Grey properties of the willow, cf. that Fi. *selja* ‘elder’ is borrowed from PGmc. *\*salhjō-* ‘willow’, which is normally regarded to be ultimately derived from PIE *\*sal-* ‘grey, dirty; ? salt-like’; the willow can grow as a parasite in elders (Tholle 1944)

As mentioned before, the linking of *Sambucus* species in popular taxonomy to *Viburnum opulus*; cf. e.g. Eng. *water elder*, NHG *Traubenholunder*, Dan. *dværghyld* and Ru. *kalína* ‘water elder’ is quite remarkable seeing that the latter has orangy berries but is still derived from the root *\*kel-* ‘black’. This must mean that its name originally designated *Sambucus nigra*; conversely, OIr. *ruis* also covers ‘*Sambucus nigra*’, although it is derived from the PIE word for ‘red’. From a designation ‘grey’ with no semantic shift we can imagine the following motivations:

- e) The particularly grey bark, common in botanic descriptions of *Sambucus* species (if used as a material)
- f) Blue-grey dye can be produced from the berries of *Sambucus nigra*

---

*vliender*, *vlerk* after *vlinderen*, *vlerken*, cf. *fladderen* and even *vlieden*, *vlien* ‘move in the wind’), almost as if by a process of phonaesthetic value of the sequence *vl-* and *fl-*, I so far hesitate to put too much weight on *vlaar*. Cf. also the parallel in Fi. *heisipuu* becoming *hörskipuu* after *hörskyä* ‘be frayed, sway’ and similar forms from Balto-Fennic below. Note that the Gmc. word is not crucial for the points presented in this paper, and that a root *\*pleu-* could theoretically also be an extension of *\*pel-*

- g) Powdery mildew on berries, canes and leaves can give elder trees a grey impression (cf. the second meaning of Lith. *šývis* ‘mildew’)

I prefer the latter possibility for two reasons. First of all, if the elder was once denoted ‘the dusty tree, the powdery tree’, we can suddenly understand the formation of its Latin name. The root *sa(m)b-* we were left with after having subtracted the productive plant-name suffix *-ūcus* from Lat. *sa(m)būcus* is identical to the root in Lat. *sab-ulum* ‘sand’ < PIE *\*b<sup>h</sup>sam-d<sup>h</sup>-o-*. I quite agree with Garnier (2006) that the causes for this alternation goes back to variants of a PIE collocation *\*b<sup>h</sup>s-m-eh<sub>2</sub>d<sup>h</sup>eh<sub>1</sub>-* ‘to reduce to powder, to pulverize’ vs. *\*b<sup>h</sup>s-ǵ<sub>2</sub>-b<sup>h</sup>o-* ‘pulverize’; cf. Skt. *bhásma-* ‘dust’ and similar meanings in other IE languages. While I believe that *sambūcus* was formed in Italic rather than all the way back to a PIE *\*b<sup>h</sup>sǵ<sub>2</sub>-m-b<sup>h</sup>-u-h<sub>3</sub>ʷo-* or PIE *\*b<sup>h</sup>sǵ<sub>2</sub>-d<sup>h</sup>-u-h<sub>3</sub>ʷo-* ‘(plant) looking dusty’, I do find it very likely, having accepted that ‘sand’ comes from a word for ‘powder, dust’, that this was still the meaning of *\*sa(m)b-* at the time of formation. This means that *sambūcus* would be exactly ‘the dusty tree’, i.e. ‘the moldy tree’, and not as such ‘the sandy tree’ which would be an anachronistic interpretation. Thus, the Latin name can correspond exactly in meaning to Lithuanian and almost with the Slavic and Germanic ones, meaning rather ‘the grey(ish) tree’, but supposedly still referring to mildew.

## 6 What loans in Fenno-Ugric can tell us

That at least the Balto-Slavic perception of the elder as ‘the grey one’ may be much older than the Slavic-Turkic contacts is indicated by the following facts :

- a) The Fenno-Permic name of the water elder is reconstructed as *\*šewe* or *\*šewa*, attested in Erza Mordvin *čev-gel* ‘water-elder berry’, and in the Permian languages : Udmurt (Votyak) *šu* ‘id.’, *šu-pu* ‘water elder’ and Komi (Zyryan) *žo(v)-pu* ‘id.’. This already appears to be a loanword (having *\*š-* and no relatives within Uralic), so I suggest that it is borrowed from a hypothesized Balto-Slavic word for ‘elder’, *\*šeĩua-* which is then only attested in Lithuanian and via this loanword. Note that the Fenno-

Permic word only denotes the water elder which eliminates possibilities b), c), d), e) and f) as naming motivations ;

- b) In Cheremis (Mari), the berries of water elder are called *šarša*, which must likewise be a loan in the first place ; in my eyes the only obvious source is East Baltic *\*šerš(t)as* ‘moldy’, cf. Lith. *šerkštas* ‘id.’ with intrusive *\*-k-*. Note once more that we are dealing with *Viburnum opulus*, this time leaving g) as the *only* motivation (since a) referred to the flowers and *\*šerš(t)as* must have meant ‘moldy’, not a specific color, at the time of borrowing.

Baltic loanwords in Balto-Fennic are quite numerous (around 330); in Saami there are some too, and Volgaic languages (Cheremis, Mordvin) have also borrowed directly from Baltic (van Pareren 2005, 2008; Suhonen 1988). However, no direct lexical exchange has so far been detected between Permic and Baltic, and at the Proto-Fenno-Permic stage (= Fennic, Saami, Volgaic + Permic), estimated around 2000 BC, Proto-Baltic proper could hardly yet have evolved as a distinct dialect of the Balto-Slavic unity.

If we accept the Slavic word as a Turkic loan, Slavic must consequently have had another word for ‘elder’ in the preceding centuries – perhaps *\*siv-* preserved in *\*šiv-* in Northern West Slavic. And since it is possible to derive the Lithuanian word from IE, it is conceivable that *\*šewa* is a manifestation of is the old Balto-Slavic term *\*šejuā*, borrowed into neighbouring Fenno-Permian languages during the Bronze Age. This would then be the first piece of evidence for a Balto-Slavic word borrowed into a Uralic stage as old as Fenno-Permian<sup>9</sup>. The etymology is confirmed by the fact that the Mari name for the berries of the same plant, *šarša*, can be traced back to another Baltic word for ‘moldy’.

## 7 The Elder Mother and Indo-European beliefs

Since both *Sambucus nigra* and *Viburnum opulus* are indigenous to all possible PIE homeland territories (including Anatolia), the Proto-Indo-

<sup>9</sup> Note that this stage will have to be considered even older by many uralicists in Finland now, since a consensus seems to have arisen that Uralic did not split into Fenno-Ugric and Samoyedic before Fenno-Ugric then split into Ugric and Fenno-Permian, but rather that Uralic was divided into West Uralic and East Uralic where the former would equal Fenno-Permic.

Europeans must have had names for them. ‘Elder’ belong to what we could justifiably characterize as “surprising lacunae in the IE lexicon” (‘sturgeon’ is another one).

The frequent renamings of both species all over Europe may be due to taboo, linked to popular beliefs of the elder being guarded by “The Elder Mother” (*Mother Hulda*, NHG *Frau Holle*, Dan. *Hyldemo(e)r* and its abilities to ward off evil; the Danish tradition is described by Tholle 1944). The Elder Mother was associated with spinning, considered a partly magical activity with links to the otherworld; this may have triggered a folk-etymological reshaping of the Balto-Fennic forms; cf. Fi. *heisi* ‘elder’, Est. *(h)õis* ‘bloom’ ← Baltic *\*žėidas* ‘bloom’) after words for ‘thread’ (see Mägiste 1970:357-358):

- a) Regular: Fi. *heisi*, *heisi-puu* ‘elder’, Est. *(h)õis* ‘bloom’
- b) Influenced by Fi. (Lönnrot) *höytö*, S Est. *hääie* ‘thread’: Fi. *höisi-*, *höysi-puu*, Olonets Karelian *höüdöi*, Lydian *höüdü-öi* ‘elder’
- c) Influenced by *hörsk(y)ä*, *hersyä* etc. ‘be frayed; sway’: Fi. *hörski-puu*, *hersi-puu* ‘elder’

In Danish folk-tradition it is told that an elder (a water elder?) will grow from the middle of Viborg Lake when the enemy (often “the Turk”) arrives, which is perhaps the reason for the English alternative name *Danewort* of the water elder; on the connection between Mother Hulda and Baltic folk-belief regarding lakes and spirits, cf. Gliwa 2005.

If these religious connotations are very old, we may even cautiously interpret ‘the grey one’ as having a double meaning, remarkably, but presumably incidentally, alike the ones covered by Mod. Eng. *elder*. In English folk-belief, Mother Elder is also known as *The Old Lady*. The Eng. name itself is reshaped from OE *ellærn*, partly after other tree-names in *-der* (cf. MLG *elderne*), corresponding to MLG *alhorn*. It may very well be that this has happened with the meaning of the homonym in mind.

## 8 Concluding remarks

We may summarize the scenario as follows: The (unknown) PIE term for ‘elder’ became replaced in NW PIE by different terms all meaning ‘the grey one’ or ‘the dusty one’, referring to powdery mildew and perhaps only later to Mother Elder. In Balto-Slavic, a reflex of *\*keih-uo-*

was used and borrowed into neighbouring Fenno-Ugric languages. In Late Common Slavic, the new indigenous term was replaced with its synonym in Old Turkic, used by adjacent Turkic or mixed Altaic tribes; perhaps the original Balto-Slavic word survived in the Northwestern-most fringes. Several of the Balto-Fennic forms were reshaped under influence from the word for ‘thread’, referring to Mother Hulda’s magical spinning techniques.

We may illustrate the course of events in a chart like this:

#### Stage 1

An unknown PIE name \*X is the original common name for two kinds of plants ‘*Sambucus*; *viburnum opulus*’, perhaps already personified as ‘mother elder’ and colloquially replaced by ‘the grey one’, ‘the dusty one’ because of taboo

>

#### Stage 2

NW PIE the term ‘the grey one’ or ‘the dusty one’ gets lexicalized and becomes the normal term for both species

>

#### Stage 3

- a) Italic uses ‘the dusty one’ (anachronistically \**b<sup>h</sup>s<sub>2</sub>-m-b<sup>h</sup>-u-h<sub>3</sub>ʷo-*)
- b) Balto-Slavic uses ‘the grey one’ or ‘the moldy one’ (\**ḱeiH-ṽo-*)
- c) Germanic uses the ‘the grey tree’ (PGmc. \**fli-þra-*)
- d) Germanic and Slavic also has ‘the black one’ (\**kel-ŋ-*) referring originally to *Sambucus nigra*, but later also to *Viburnum opulus*
- e) Celtic may have used ‘the red one’ (OIr *ruis*) of certain species and later transferred it to *Sambucus nigra*

>

#### Stage 4

The Balto-Slavic word is borrowed to Proto-Fenno-Permic (\**šewa*) as the designation for *Viburnum opulus* only

>

#### Stage 5

Another (East) Baltic expression for ‘moldy’ (\**šerša-*) is borrowed by Mari (Cheremis) as the name for the berries of *Viburnum opulus*

>

#### Stage 6

- a) Late Common Slavic replaces the domestic word with a Turkic synonym \**buzv* > *buz*; the original word survives in the extreme Northwest as \**šiv-* (Sorbian, and as a loan in German dialects)

- b) Balto-Fennic forms are reshaped after names for 'thread' (*heisi* 'elder' > *höisi* after *höytö* etc.
- c) The Old English form *ellærn* is reshaped after other tree names in *-der* (such as *alder*), probably not creating a homonym for 'old person' by sheer coincidence.

## References

- Andersen, Henning, "Slavic and the Indo-European Migrations". – Henning Andersen (ed.): *Language Contacts in Prehistory: Studies in Stratigraphy*. Amsterdam/ Philadelphia: John Benjamins. Pp. 45-76.
- Birnbaum, Henrik, 1998: "Common Slavic in Time and Space". – *Scando-Slavica* 44:131-143.
- Blazek, Václav, 2002: "The "Beech"-argument: State-of-the-Art". – *Historische Sprachforschung* 115: 190-217.
- Brüch, Joseph, 1922: "Lateinische Etymologien [4. Lat. *sabucus*, *sambucus*]". – *Indogermanische Forschungen* XL: 232-241.
- Dyneley Prince, John, 1919: "Tatar Material in Old Russian". – *Proceedings of the American Philosophical Society* 58: 74-88.
- Garnier, Romain, 2006 [2008]: "Nouvelles reflexions autour de Grec ψάμαθος". – *Die Sprache* 46/1: 81-93.
- Gliwa, Bernd, 2005: "Einige litauische Ortssagen. Bodenlose Gewässer und Frau Holle (KHM24, ATU 480)". – *Studia Mythologica Slavica* VIII: 187-223.
- Gliwa, Bernd, 2008: "Litauisch *šeivamedis*, *šeivą trinti*, Lettisch *saiva*, Finnish *käävi*". *Baltu Filologija* XVII (1/2): 41-60.
- Henning, Walter B., 1963: "The Kurdish Elm". – *Asia Major* 10: 68-72.
- Kluge-Seebold 1989 = Friedrich Kluge: *Etymologisches Wörterbuch der deutschen Sprache*, 22. ed. revised by Elmar Seebold. Berlin / New York: Walter de Gruyter 1989.
- Mägiste, Julius, 1959: "Gibt es im Tscheremissischen baltische Lehnwörter?". – *Ural-Altaische Jahrbücher* 31: 169-176.
- Mägiste, Julius, 1970: "Estn. (*h*)õis 'Blüte' und seine Sippe – ein Balt. Lehnwort?". – Velta Rūķe-Draviņa (ed.), *Donum Balticum* [Fs. Stang]. Stockholm: Almqvist & Wiksell. Pp. 352-360.
- Mucke, Karl Ernst, 1891: *Historische und vergleichende Laut- und Formenlehre der niedersorbischen (niederlausitzisch-wendischen) Sprache mit besonderer Berücksichtigung der Grenzdialekte und des obersorbischen*. Leipzig [reprinted by Zentral-Antiquariat der DDR 1965].
- van Pareren, Remco, 2005: *Baltische leenwoorden in het Mordwiens*. MA thesis from the University of Groningen.
- van Pareren Remco, 2008: "Die direkten baltischen Lehnwörter im Mordwinischen". – *Finnish-Ugrische Mitteilungen* 30-31: 69-147.
- Räsänen, Martti, 1969: *Versuch eines etymologischen Wörterbuchs der Türkssprachen* [= *Lexica Societatis Fenno-Ugricae* XVII]. Helsinki: Suomalais-Ugrilainen Seura.

- Schwyzer, Eduard, 1934: "Dissimilatorische Geminatenauflösung als Folge von Übersteigerung usw.". – *Zeitschrift für Vergleichende Sprachforschung* 61, 3/4: 222-252.
- Suhonen, Seppo, 1988: "Die baltischen Lehnwörter der finnisch-ugrischen Sprachen". – Denis Sinor (ed.): *The Uralic Languages: Description, History, and Foreign Influences*. Leiden / New York / Copenhagen / Cologne: E.J. Brill. Pp. 596-615.
- Tholle, Johannes, 1944: "Hyldebusken". – *Danske Studier* 1944: 1-38
- Vasmer, Max, 1953-1958: *Russisches etymologisches Wörterbuch*. Heidelberg: Winter.
- Wichmann, Yrjö, 1903: *Die tschuwassischen Lehnwörter in den permischen Sprachen* [= *Mémoires de la Société Finno-Ougrienne* XXI]. Helsinki: Suomalais-Ugrilainen Seura. 98-99
- Wichmann, Yrjö (original manuscript), T.E. Uotila & Mikko Korhonen (eds.), 1987: *Wotjakischer Wortschatz* [= *Lexica Societatis Fenno-Ugricae* XXI]. Helsinki: Suomalais-Ugrilainen Seura. 231.





## Balto-Fennic *\*kakra* ‘oats’, the Etymology of *hail*, and another Exception to the Germanic Sound Shift

### Abstract

Labials develop regularly into velars when preceding *-l-* and *-N-* in various Scandinavian dialects past and present. Since delabialization is not common before *-r-*, Nw. dial. *hagre*, OGutn. *hagri*, Mod.Gutn. *hagre*, Elfd. *ager*, all ‘oats’, have been interpreted as reflecting an original PGmc. variant *\*hagran-* alongside the more widespread *\*habran-*. Kroonen (2013) separates the two forms, reconstructing a meaning ‘broom grass’ for *\*hagran-*, with reference to especially Da. *hejre* id. However, it has been overlooked that Proto-Balto-Fennic *\*kakra*, an early Germanic loan, points to a meaning ‘oats’ attached to the velar variants in Proto-Germanic already. Witczak’s (2003, 2004) PIE reconstruction *\*k̑op-r̥* ‘grass; vegetables’ is accepted, but a semantic and derivational stratigraphy reveals that core IE remnants of the old heteroclitc stem (such as Alanic *zabar*) acquired a specific meaning ‘oats’, while purely thematicized forms were attached to less specific, non-agricultural, meanings. PGmc. *\*hagla-* ‘hail’ can be related; deriving terms for ‘hail’ from ‘grain’ is typologically common, and a sound-law whereby PIE *\*-opl-*, *\*-apl-* yields *\*-agl-* or *\*-ak(u)l-* would allow for five other Germanic words of disputed origin to be assigned quite straightforward etymologies: *\*tagla-* ‘tail’ could go back not to *\*dōk-lo-* but to *\*dop-lo-*, derived from the root in Sw. *tafse* ‘tuft’—an alternative also applicable to OIr. *dúal*, but leaving out otherwise irregular Slavic forms; OSw. *sakla* ‘to drool’ is linked to EFris. *sabben* ‘to drool’ and PGmc. *\*safta-* ‘juice’ rather than NWGmc. *\*sakkan-* ‘sink slowly’, Sl. *\*sokŭ* ‘juice’; Da. *rakle* ‘catkin’, Sw. dial. *rackel* ‘long thing; tall, slim person’ are equated with Lith. *rāplės* ‘thongs’, akin to Sl. *\*repĭji* ‘burdock’, Alb. *rrap* ‘plane tree’, PGmc. *\*raſtra-* ‘long, thin pole’; *\*skakulaz* ‘whippletree; traces for harness horse; schackle’ is deemed identical to Lat. *capulus* ‘halter for horse; towing rope; handle’, *scapula* ‘shoulder’, derived from

the root in *scāpus*, PGmc. *\*skaftaz* ‘shaft’; and *\*hakulaz* ‘cloak’ is seen as a derivative of the root in MLat. *cappa*, *cāpa* id., which can be shown to be of IE rather than Semitic origin.

### 1 Balto-Fennic *\*kakra* and Germanic words for ‘oats’

Labials have developed into velars in the position before *-n-* and *-l-* in various North Germanic dialects, cf. for *\*-pn-* ON *vákn* ‘weapon’, OSw. (Västgötaland, Östgötaland, Hälsingland, Västmanland, Söderkalix) *vākn*, Far. dial. *vákn* and even Fi. *vaakuna*, an old Scandinavian loanword, next to ON *vápn*, OSw. *vāpn*, Far. *vápn* < PGmc. *\*wēpna-* id.; Norrland Sw. *göcken*, Mod.Gutn. *gauken*, Far. *geykn* ‘handful’ ~ Sw. *göpen*, MDa. *giøben*, ON *gaupn* ‘both hands held together’, OHG *goufana* ‘handful; empty hand, palm’ < PGmc. *\*gaupnō-*; OSw. (Västgötaland) and Norrland Sw *strågna* ‘suffocate’ < *\*stropna*, from *strypa* ‘strangle’; for *\*-fn-* OSw. *ughe*n, Sw. *ugn*, SW Nw. and Inner Trøndelag Nw. *ogn*, Icel. *ónn* ‘oven’ next to ODa. (Haderslev Stadsret 1292) *ufen*, Sw. dial., Nw. *omn* < PGmc. *\*ufna-* id.; for *\*-pl-* Nw. *drøygla* ‘to secrete from the uterus (of a cow in rut)’ ~ Nw. *dropla* ‘drip quietly’, PGmc. *\*draupa-* ‘a drop’; and for *\*-fl-*, *\*-bl-* *\*-vl-* Sw. *dregla* ‘to drool’ ~ OE *dreflian* ‘to dribble or run at the nose, slobber’, Eng. *drivel*, *drool* < PGmc. *\*drablōjan*-<sup>2</sup>, OSw. *sughl* ‘meat’ ~ OSw. *sufl*, Sw. *sovel* < PGmc. *\*subla-*; and OSw. *swaghel*, Sw. dial. *svagel* ‘sulphur’ ~ Standard Swedish *svavel*, borrowed from MLG *swavel*.

<sup>1</sup> For delabialization in this item in Nordic specifically, cf. Bjorvand & Lindeman (2001: 704) and Kroonen (2013: 446). Traditionally, OSw. *ughe*n, Sw. *ugn* have been considered Verner variants of another form *\*uh(w)na-* reflected in Goth. *aúhns*. However, Gothic in fact seems to have undergone a similar, although more restricted, development of *f* > *h* before nasal (cf. *aúhuma* ‘upper’ < *\*ufuman-*), not contested by any obvious counterexamples (cf. also *lauhmuni* ‘lightning’ in fn. 40; see Hyllested & Cohen 2007: 15–16). Not a single Germanic attestation thus points unambiguously to a variant with labiovelar *\*uh(w)na-* or *\*ug(w)na-*; hence, we need not reconstruct any other forms for ‘oven’ in Germanic than *\*ufna-*, with perfectly regular developments throughout the daughter-languages (that is, if one disregards late dialectal confusion resulting in variation in, say, Modern Standard Swedish where *†umn* would be the expected outcome).

<sup>2</sup> These etymologies for Sw. *dregla* and Nw. *drøygla* respectively are new. The two words have otherwise been grouped together (Tamm & Noreen 98; Hellquist 1898) but are usually not assigned further certain etymology.

Mostly, this is described as a sporadic, yet common phenomenon<sup>3</sup>, However, judging from the material in e.g. Rydquist (1868: IV, 239, 249), it appears quite if not altogether regular in some dialects of (Old) Swedish (especially Norrland and Västgötaland) and Norwegian (confined to Old W Norwegian, Mod. SW Norwegian and the dialects of Inner Trøndelagen [Lierne]), and it dominates in Faroese<sup>4</sup>, while, on the other hand, it is completely absent from (Middle) Danish<sup>5</sup> and Old Gutnish<sup>6</sup>.

<sup>3</sup> Thus Wessén 1965 (I: 46) for Swedish in general, Olson (1904: 116) for the old dialects of Östgötaland and Seip (1931: 188) for Old (S)W Norwegian. However they only mention the development of *\*p > k, g/\_n*.

<sup>4</sup> Also after loss of intervening *-t-*, cf. *akna* ‘become evening’ < ON *aftna* id.; Older Far. <*okn*> (= *ákn*) ‘swan’ < *\*álptn-* (Lockwood 1961: 57).

<sup>5</sup> Unfortunately the situation in Modern Danish is often ambiguous because of the merger between *\*-g-* and *\*-v-* and sometimes *\*-b-* in relevant positions, and the possibility of spelling the ultimate result [w] both with <g> and <v>, and sometimes even <b>. This is also the case for forms with <v> in Bokmål Norwegian, even pronounced with labiodental [v], that are historically Danish (e.g. *ovn* ‘oven’).

<sup>6</sup> According to Kroonen (2013), OSw. *ljung-eld*, Mod.Gutn. *liaugn* ‘lightning’ constitute another example and must come from Old East Norse *\*ljūfn* (OSw. *lyghna* f. id. < *\*-jōn*), which would in itself be a secondary labialization of Proto-Norse *\*-uhn-* into *\*-ufn-*. Kroonen here implies that an old Verner variant PIE *\*leuk-nó-* would not have yielded PGmc. *\*leugna-* but *\*leukka-* with Kluge’s Law instead. Nw. *ljon*, *lyn*, Da. *lyn* ‘lightning’ (and Mod.Icel. *ljón* ‘hurricane’, attested from the 19<sup>th</sup>.c.) would still point more directly to a Germanic and Norse protoform with *\*-hn-*. At first glance, the supposition that *\*ljūfn* was a general East Norse form seems to be contradicted by ODan. *liughnæth* ‘lightning’ since the developments *\*-f/v- > -g- /\_l/n* and *\*-p- > -k- /\_l/n* do not occur in Old Danish; however, *-gh-* in this word can simply reflect a difficulty of identifying the fricative [ɣ] following *-u-*, cf. the orthographic variants *frughæ ~ fruæ* << OSax. *frúa* or MLG *früwa* and (Brøndum-Nielsen). On the other hand, *liughnæth* is the only example where a purely “orthographic” (unetymological) *-gh-* would occur outside hiatus, so we cannot exclude that *\*-ghn-* does reflect a Proto-Norse *\*-gn-*, which, because of Kluge’s Law, must still be explained as secondary from a Proto-Germanic point of view. In fact, no Germanic forms even point unambiguously to its unvoiced counterpart, PGmc. *\*leuhna-*. In Gothic, first of all, *láuhmuni* f. ‘flash’ may straightforwardly reflect *\*laufmunjō-* (see the previous footnote). Second, we cannot be completely sure of the regular outcome of a PGmc. full-grade sequence *\*-eufN-*, *\*eubN-*, *\*-aufN-*, or *\*-aubN-* in Old West Norse; what we have apart from Nw. *ljon* ‘lightning’, Mod.Icel. *ljón* ‘hurricane’ (attested only from the 19<sup>th</sup>.c.) is the variation between ON, Mod.Icel. *ofn* ‘oven’ on one hand and Mod.Icel. *ónn* ‘...’, attested from the 1500’s, on the other. There is a possibility that *ónn* is borrowed from Danish (since *ó* from that century onwards has been pronounced as a diphthong close to Da. *-ov-*), but ÍO prefers an old variation in OVN. The latter remaining a pos-

Since delabialization is not common before *-r-* anywhere, Nw. dial. *hagre*, OGutn. *hagri*, Mod.Gutn. *hagre*, Elfd. *ager*, all ‘oats’, have been interpreted as old forms reflecting a PGmc. variant *\*hagran-* (~ *\*hagrōn-*) m. alongside the more widespread *\*habran-* (~ *\*habrōn-*) m. (> ON *hafri*, Sw. *havre*, OS *haboro*, OHG *habaro*, NHG *Hafer*)<sup>7</sup>. While the forms reflecting a variant with velar are simply left out in Orel (2003), Kroonen (2013) tries to separate the two variants historically, and on the basis of especially Da. *hejre* ‘brome grass’ and, with less weight, Mod.Icel. *hellin-hagra* ‘mother-of-thyme, *Thymus praecox*; wild thyme, *Thymus serpyllum*’<sup>8</sup>, he tentatively reconstructs the meaning ‘bromus’ for PGmc. *\*hagran-*<sup>9</sup>. He further notes “Brome grass is closely related to the wheat-grass lineage, and is known to infest grain fields”, leaving open various possibilities of relationship to PCelt. *\*korkio-* ‘oats’, OIr *corc* ‘hair’ and Nw. dial. *hagr*, *harg* ‘coarse hair from a horse’s mane or tail’ via dissimilation or metathesis<sup>10</sup>. In any event, he chooses to separate *\*hagran-* with a velar from *\*habran-* with a labial, assuming an old meaning ‘oats’ only for the latter.

Both Orel and Kroonen derive *\*habran-* from *\*habra-* ‘goat’, thus operating with an inner-Germanic origin of this word. However, Witczak (2004) makes an excellent case for a broader Indo-European

---

sibility, we cannot exclude that Icel. *ljón*, Nw. *ljon* reflects the same kind of variation, and that *\*-eufN-* and *\*-euhN-* merged in (variants of) OWN. The remaining form mentioned by Kroonen, ME *lēven* ‘lightning, flash, flame’, is traced back to a form *\*lauhumna-* but may equally well reflect *\*laufumna-*. Now, even a lone PGmc. *\*leuf-* needs an explanation: From an Indo-European point of view, we would obviously expect ‘lightning’ to be derived from *\*leuk-* ‘light’, which occurs in other derivatives as PGmc. *\*leuh-*; however, the word for ‘lightning’ in ON is actually *leiptr* (Mod.Icel. *leiftur*) which belongs with Lith. *liepsnà* ‘id.’, going back to an altogether different root *\*leip-*. A Proto-Germanic contamination between *\*leuh-* in *leuhman-* ‘beam of light’ (ON *ljómi*, OE *lēoma*) etc. on one side and *\*leif-* in *leiftraz* ‘lightning’ (ON *leiptr-*) on the other, resulting in new PGmc. words for ‘lightning’ beginning with *\*leuf-*, appears to have been unavoidable.

<sup>7</sup> Thus, Da. *havre* can theoretically reflect both *\*habra-* and *\*hagra-*.

<sup>8</sup> Glossed in Kroonen (2013) as ‘a kind of thyme’. The common name in Modern Icelandic today is *blóðberg*.

<sup>9</sup> Elsewhere in the entry he supposes a “broader meaning” than just ‘oats’.

<sup>10</sup> An equestrian term common to Celtic and Germanic would match the results reached in Hyllested 2010 quite well; there, I argue that old vocabulary common to Germanic and Celtic can be grouped into a few semantic categories; one of the important subcategories is exactly terms pertaining to horsebreeding and riding. Another word meaning ‘horsehair’, shared by Celtic and Germanic, is mentioned later in this article.

word family involving Hitt. *kappar-* ‘vegetables’; Skt. *śāpa-* m. ‘drifting reed; reed washed onto a riverbank’; Alan. *zabar* ‘oats’, ModPers. *sabz* ‘grass; green; dark’, Pashto *sābah* ‘grass; vegetables’, Shughni *sip(i)yak* ‘a kind of millet’ (< PIr. *\*sāpar-ku-*), Roshani *sabēc* ‘pod of bean’; Lith. *šāpas* ‘straw, grass’, *šāpai* ‘what remains in the field after a flood’; MIr. *corca*, *coirce* ‘oats’; W *ceirch*, MBret. *querch* (in the Celtic forms, *\*-p-* has regularly disappeared, and *-ch-* is seen as reflecting a suffixal *\*-k-jo-*, cf. the Iranian forms). Although Witczak reconstructs both PIE forms, a heteroclitic *\*kōp-r-* and a thematic *\*kōp-o-*, with very broad meanings ‘grass; vegetables’, especially the Alanic meaning underlines that the meaning ‘oats’ can be old. More specifically, ‘oats’ seems to have become the specialized meaning of the non-Anatolian IE remnants of the heteroclitic form, while the thematic form is rather attached to meanings like ‘grass’, ‘straw’, or dead plants floating in the water. In PIE including Anatolian, the exact meaning is less precise, but it appears that PGmc. n-stems are built upon secondary thematicizations to the original heteroclitic.

Several other of Kroonen’s points can be contested. Designations for ‘brome grass’ can be transferred from those from ‘oats’ or vice versa simply because of the strong physical similarity between the plants; cf. e.g. Danish *hejre-havre* ‘bromus’. An Icelandic name like *hellin-hagra* does not tell us much about the original meaning of *\*hagrōn-* since it only makes up the last part of a compound of a type common for grass-like or edible plants or fodder, cf. besides *hejre-havre* also Dan. *drap-havre* ‘oatgrass, Arrhenaterum’. Furthermore, *hejre* could theoretically be connected to Nw. *hagr*, *harg* ‘coarse hair from a horse’s mane or tail’ (and Shetland Norn *hegri* ‘thin yarn of wool’) without having anything to do with the ‘oats’-word, cf. the almost exact parallel source of Nw. *fakse*, Sw. *faxe* ‘bromus’ < ON *faxi* ‘horsetail’. While I do find it possible that a contamination or maybe even a reshaping took place involving the horsetail-word, so that for example a PGmc. *\*harga-* ‘horsehair’ merged with *\*habra-* ‘oats’ into the variant *\*hagra-* ‘oats’, this would not have taken place earlier than in Common Scandinavian for the following reasons:

The velar variant is attested in all languages of the Balto-Fennic group, everywhere with the meaning ‘oats’, and unanimously pointing to a Proto-Balto-Fennic form *\*kakra*, cf. Fi. *kaura*, E dial. *kakra*, Karel. *kagra*, Veps *kagr*, Est. *kaer*, Vot. *kagra*, Liv. *kagr*. Since Proto-Balto-Fennic is usually regarded to be at least as old Proto-Germanic, we must maintain a form *\*hagra-* with the meaning ‘oats’ alongside the

synonymous *\*habra-* already in Proto-Germanic (although it is of course theoretically possible that the various branches of Balto-Fennic borrowed the word individually at a later stage, say, from a stage close to Proto-Norse, a possibility not preferred by Ockham's razor). In other words, Balto-Fennic confirms that *\*hagra-* did not primarily mean 'bromus' which would perhaps be a less obvious culture-word anyway.<sup>11</sup>

Summing up, it seems clear that PGmc. *\*hagran-* m. 'oats'<sup>12</sup> and *\*habran-* m. both existed in Proto-Germanic as variants of a word meaning 'oats', and that the latter variant was borrowed into Proto-Balto-Fennic as *\*kakra*. It is an n-stem formed from a secondary thematicization *\*kōp-ró-* of an old heteroclitic *\*kōp-r̥* which in core IE acquired a meaning 'oats' (in some Eastern Iranian languages, it has further come to signify a local kind of millet). Derivatives outside the heteroclitic system retained meanings of a less specific and less agricultural character such as 'grass', 'vegetables', 'straw' or 'plants flowing in water'. This provides us with a stratification of meanings that neatly match what we claim to know about the cultural history involved: That the speakers of PIE were less hard-core agriculturalists than the speakers of later Western IE offshoots.

## 2 PGmc. *\*hagra-* 'oats' and *\*hagla-* 'hail'

Regardless of its precise meaning, the existence of a PIE root *\*k̑ap-* or *\*k̑ep-* used in words for crops, everywhere occurring in what could be

<sup>11</sup> Neither do I agree with Hofstra (1995: 95) that Fi. *kakra* and its Balto-Fennic cognates can be loanwords from Old Gutnish during the middle ages. As mentioned before, the delabialization is not even specifically Gutnish in the first place, cf. *vapn* 'weapon', *hamn* 'harbor', *stefna*, *stemna* 'to point out', *afla* 'breed', *gafl* 'fork', *cafli* 'piece cut out', although at least *gauken* 'handful' and probably *liaugn* 'lightning' show that it occurs (sporadically) in Modern Gutnish which is not an uncontaminated descendant of Old Gutnish. What we need is a the most plausible scenario to explain the presence of velar forms in all Balto-Fennic languages.

<sup>12</sup> Several people have contested my use of the term "delabialization" in this case since by delabializing a labial you do not get a velar. I maintain the use of the term, for which there is no good alternative, because delabialization in this case refers to phonological rather than phonetic conditions. The velar stop is simply the stop you get by removing the feature [+lab] from the labials. Typologically, dentals also become velar when preceding sonorants, so that velars can be said to make up the unmarked stops in this position.

the o-grade, paves the way for etymologizing the otherwise mysterious PGmc. *\*hagla-* ‘hail’.

According to the etymological standard handbooks, PGmc. *\*haglaz* m. ~ *\*haglan-* n. ‘hail’ (> ON *hagl*, OE *hæz(e)l*, *hazol*, OFris. *heil*, OS, OHG *hagal*) derives from a PIE protoform *\*kagh-lo-*, whence also Gk. *κάχληξ* ‘pebbles’. The Germanic-Greek equation indeed appears acceptable from both a semantical and a formal point of view. However, Greek also possesses unexplained variants like *κόχλαξ* with divergent vocalism, and *ἄχλαξ* where the first consonant is missing; such a variation is normally interpreted as a sign of Pre-Greek substratum origins (Beekes 2010 : 606 ; Kroonen 2013)<sup>13</sup>. Bjorvand & Lindeman (2000: 336), as well as Orel (2003: 150) believe that we are rather dealing with an inherited Verner variant of a PIE o-grade noun, *\*kókló-*, whereby *\*hagla-* corresponds to PGmc. *\*hahla-* ‘slippery’ with original stress on the first syllable, *\*kóklo-*. This word would have been formed by means of reduplication from the root *\*kelH-* ‘freeze’ (reconstructed with a laryngeal because of the acute in Lith. *šálti* ‘freeze (of ice)’); cf. also *pā-šalas* ‘frozen ground’, Du. *hal* id.; reduplicated formations with *-i-* in the first syllable are known from Skt. *śíśira-* ‘cold’, ON *hēla* ‘hoarfrost’ < *\*ki-klH-o-*. Neither etymology can be excluded, and, so far, the origin of *\*hagla-* must be regarded as unsolved.

Semantic typology often proves a fruitful starting-point in etymology: When designations for ‘hail’ actually are synchronically analyzable, what do they then reveal? In many languages they turn out to constitute parallels to the classical etymology involving the Greek word for ‘pebbles’; beginning with Modern English, a single piece of hail is called *hailstone*, in Dutch correspondingly *hagelsteen* and in Portuguese *pedra de granizo*, literally ‘stone of hail’. This is also the most common way of designating hailstones in the older Germanic languages: ON *haglsteinn*, OE *hazolstān*, MLG *hagelstēn*, MHG *hagelstein*. On the other hand, calling hailstones ‘grains of hail’ appears typologically just as common, cf. e.g. NHG *Hagelkorn*, Da. *haglkorn*, It. *chicco di grandine* and the etymology of Sp. *granizo* (a derivative of Lat. *grānum* ‘grain’ which is also the source of the last part of the Portuguese word). Besides, the term for the hail-like meteorological phenomenon ‘graupe’ contain words for ‘grain’ in Da./Nw. *iskorn* and Dutch *ijskorrel*, lit. ‘ice-grain’, and the word *graupe* itself is related to both MHG *īs-grūpe* ‘hail’ and NHG *Graupe* ‘grain’, historically a derivative of PGmc. *\*greupan-* ~ *\*grūpan-* ‘to cut

<sup>13</sup> It should be noted that Beekes represents the Leiden school where *a*-vocalism in itself is regarded as a typically non-Indo-European feature.



up, grind' (Kroonen 2013). Furthermore, in Danish and Norwegian lexicography, hail itself is typically defined as *iskorn*, literally 'grains of ice'<sup>14</sup>.

The connection between hail and grain goes further: the H-rune in the runic alphabets, which in the younger futhark looks much like a stylized snowflake, was called *hagall* in Old Norse and *hægl* in Old English. Although these names are not completely identical to the word for 'hail' (ON *hagl* < \**hagla*- n., not *hagall* < \**haglaz* m.) there is no doubt that the name refers to hailstones. In the Anglo-Saxon rune poem from the 8th or 9th century, it is described in the following way :

*Hægl byþ hwītust corna*  
*hwryft hit of heofnes lyfte*  
*wealcop hit windes scūra*  
*worþeþ hit tō wæþere syððan*

Hail is the whitest of grains ;  
 It whirls down from the heaven's height (air)  
 And gusts of wind toss it about;  
 Then it is transformed to water

The Norwegian rune poem, dating back to the 13th. century, reads:

*Hagall er kaldastr korna*  
*Kristr skóp hæimenn forna*

Hail is the coldest of grain;  
 Christ created the world of old

The Icelandic rune poem, from the late 15<sup>th</sup> century onwards, says:

*Hagall er kaldakorn*  
*ok knappa drífa*  
*ok snáka sótt*

Hail : cold grain

<sup>14</sup> In the meaning 'graupel' it is less of an everyday word and therefore sufficiently little lexicalized to be used as a productively formed noun whose two parts are understood literally.

and driving sleet  
and sickness of snakes<sup>15</sup>

Griffiths (2006) argues that in the famous Old Irish work *Auraicept na n-Éces* (‘the scholar’s primer’, originally from 650, edited until around 1150, preserved in the *Book of Ballymote* from 1390), a word-play involving the word for ‘grain’, OIr. *gráin*, can lay behind the naming of the h-rune in the Ogham alphabet, (*h*)*Úath*, since both *húath* and *gráin* also mean ‘horror’. The word-play would then also involve Latin *horror*, meaning both ‘horror’ and ‘shivering (from cold)’. In the so-called arboreal tradition, where Ogham runes are named after trees, (*h*)*Úath* is glossed as ‘white-thorn’ (probably referring to hawthorn, but according to Griffiths deliberately described like hail). Judging from the different kennings of the rune, it is obvious that the rune-name is connected metaphorically and mythologically with its homonym *húath* ‘horror’ (translations by McManus 1988):

1 (Bríatharogam Morainn mac Maoín)  
*condál cúan*  
‘assembly of packs of hounds’

2 (Bríatharogam Maic ind Óc)  
*bánad gnúise*  
‘blanching of faces’

3 (Bríatharogam Con Culainn)  
*ansam aidche*  
‘most difficult at night’

Thus, the typological perception of hail as ‘grain’ is clear. It is therefore worth considering if PGmc. \**hagla-* contains the same root as \**hagra-* ‘oats’; cf. that *corn* in Scots and Hiberno-English means both ‘oats’ and ‘grain’ in general. Such conditions could also explain the similarity between the most widespread Indo-European words for ‘hail’ and ‘grain’, respectively: a) Lat. *grandō*, OCS *gradŭ*, Arm. *karkowt*; and b) Lat. *grānum*, PGmc. \**kurna-*. Lat. *grandō* results either from regular metathesis *-nd-* < \**-d-n-* following suffixation with \**-no-* (Rasmussen 1999 [1984]: 152-154) or a contamination of *-d-* in the ‘hail’-word with *-n-* in

<sup>15</sup> Translations from Griffiths 2006: 90.

the ‘grain’-word. Cf. that although Spanish and Portuguese derive from Latin, their words for ‘hail(stone)’, despite the great similarity, do not derive regularly from the Latin ‘hail’-word, but seem to be derivatives of the ‘grain’-word.

More than anything, *\*hagla-* now simply looks as if it contains the same root as our ‘oats’-word, obviously with a similar delabialization like the one in *\*hagran-*. This kind of delabialization is already known to occur sporadically in Germanic preceding liquids, cf. e.g. Dutch *heuvel* next to NHG *Hügel* ‘hill’, and Limburgish *swegel* < *\*swebla-* ‘sulfur’ (Kroonen 2013).

Now, one could of course hypothesize that *\*-g-* for expected *\*-b-* not only in *\*hagra-* but even in *\*hagla-* is due to sporadic delabialization or word contamination. However, for the latter it seems more likely that a sound-law is applying. On closer inspection, it indeed turns out that examples with a preserved labial before *-l-* and following PGmc. *\*-a-* are totally absent: There are simply no cases of the expected outcome PGmc. *\*-afl-* ~ *\*-abl-* from PIE *\*-apl-* or *\*-opl-*. Quite a few instances of Germanic *\*-afl-*, *\*-abl-* do occur, but they always come from PIE *\*-ab<sup>h</sup>l-*, *\*-ob<sup>h</sup>l-* (e.g. PGmc. *\*kablōn* m. ‘a piece cut off’ > ON *kafli*, OLFr. *cavele*, cf. Lith. *žābas* ‘branch’)<sup>16</sup>. This indicates that *\*hagla-* could be the regular

<sup>16</sup> Since I first made this claim, Kroonen (2013) has presented two candidates in his dictionary that seem to contradict it: *\*afla-* ‘hearth’ (ON *afl* id., Far. *alvur*, *alvi* ‘fireplace, forge’) which he compares to Hitt. *ḫappenaš* ‘baking-kiln, fire pit, broiler (oven)’, tentatively reconstructing an *l/n*-stem *\*h<sub>2</sub>ép-(ō)l*, gen. *\*h<sub>2</sub>p-n-ós*. He admits that the Hitt. word can alternatively be compared to PGmc. *\*ufna-* ‘oven’, which, according to him looks like an old *wanderwort*. Even if one is willing to accept the existence of PIE *l/n*-heteroclitics different from *r/n*-heteroclitics (on this topic, see now Kerkhof 2012), *\*ufna-* has a better match in Hitt. *ḫuppar* ‘oven’ (< PIE *\*h<sub>2</sub>up-r-*, *\*h<sub>2</sub>up-n-*), while Hitt. *ḫappenaš* rather belongs with Gk. *ὀπτός* ‘baked’ (< *\*h<sub>3</sub>ep-*). Original PIE word-final *-l* probably already developed into *\*-r* while the few safe examples seem to reflect a conditioned preservation of *\*-l* in *\*-uōl-/uēn-* stems only. In this particular case, one might admittedly argue that *-ōl-* could represent original *\*-uōl-* whose labial consonant would automatically be lost after *-p* (as would be the case after any other labial); this demands, however, that the loss of labial is older than the development of *-l* > *-r* which does not seem likely. Kroonen’s other example is the homonymous *\*afla-* ‘strength, power’ (ON *afl* n., Far. *alv* n. id.), reconstructed as *\*h<sub>3</sub>ep-lo-* on the basis of Hitt. *ḫapp-*. The Germanic noun is traditionally reconstructed as *\*afalan-* (OE *afol*) where the labial does not immediately precede the *-l-*. The PIE reconstruction in *-elo-* (classical *\*apelo-* before the inclusion of Anatolian reflexes) is based mainly on Greek which admittedly proves little since Greeks apparently avoids even original suffixation with *-l-* directly added to a root, replacing it with *\*-elo-*. Even from a purely Germanic viewpoint, a

reflex of *\*k<sup>h</sup>op-lo-* or *\*k<sup>h</sup>ap-lo-*, possibly in PIE already, but more likely in a late stage of (non-Anatolian) IE when more agricultural meanings of the original heteroclitlic had arisen.

### 3 Other etymologies with delabialization

Regular delabialization of PIE labials into velars between a PGmc. low vowel on one side and *-l-* on the other would render it possible to clarify five additional, otherwise disputed etymologies:

#### 3.1 PGMC. *TAGLA-* ‘(HORSE)TAIL’

PGmc. *\*tagla-* ‘horsetail’ (> ON *tagl*, Goth. *tagl* id.; Da. *tavl*, Zealand dial. *taggel* ‘horsetail’; Nw. dial. *tagl* ‘fibre’; OE *tægl*, Mod.Eng. *tail*, OHG *zagal* ‘tail’) corresponds to OIr. *dúal* ‘tail’ and is normally reconstructed as PIE *\*do<sup>h</sup>k-lo-*. Whether *-k<sup>h</sup>-* here was palatal or not cannot be decided on the basis of Germanic and Celtic alone; at first glance, a plain velar seems to be needed to account for the alleged Slavic cognate *\*dolka*, yielding SCr. *dlàka* ‘a single hair’ and Cz. *dlák* ‘branch’. Matasović (2009: 102) reconstructs a palatal to be able to include Skt. *daśā-* (f.) ‘fringe’<sup>17</sup>. However, this word-family rests on a shaky ground; notwithstanding the semantic developments, in Slavic an irregular metathesis is needed, involving the original suffix (with subsequent depalatalization of the original palatal before a sonorant), and, besides, it is far from certain that the Common Slavic form was *\*dolka* since there are no descendants outside Serbo-Croatian and Czech where *\*-olC-* og *\*-ūlaC-* merge into *-láC-*. Another possible Slavic reconstruction is thus *\*dūlaka*. From a semantic viewpoint it is just as possible that the Slavic words are related to ON *tálkn* ‘baleen’ ~ MLG *tolle* ‘branch’, from PIE *\*del-g<sup>(h)</sup>-*. OIr. *-úaR-* can come from *\*-akR-*, but is also regular from *\*-apR-*, cf. *cúan* ‘harbor’ next to PGmc. *\*hafna-*, meaning the Old Irish form *dúal* is fully compatible also with an original PIE *\*-p-* instead of *\*-k-*.

Reconstructing *\*-p-* i *\*tagla-* furthermore has the advantage that it renders possible an equation with other Germanic words of similar

---

vowel seems to have intervened, thus preventing the delabialization from taking place at the time in question.

<sup>17</sup> Matasović (2009) reconstructs *\*do<sup>h</sup>k-eh<sub>2</sub>* but the lack of Brugmann’s Law dictates e-grade.

meaning, but without the *-l*-suffix: Dan. *tave* ‘fibre’, *tavse*, *tjavs* ‘tuft’, Sw. dial. *tafse*, S Sw. dial. *tav(e)* ‘tuft’ and even PGmc. *\*tappan-* ‘tap’. A close connection between these forms and *\*tagla-* may be reflected not only in the meaning of Nw.dial. *tagl* ‘fibre’, but also in the Scandinavian sayings Nw. *med topp og tagl*<sup>18</sup>, Da. *med top og tavl*, *med tap og tavl* ‘completely (i.e. with all its body)’. These sayings of course function as alliterational figures regardless of the etymologies of their elements, but the Danish expressions with *tap* instead of *top* at least show that they can be hendiadytic (lit. ‘with hair and hair’ = with every [kind of] hair) rather than referring to two opposite extremities like the type Eng. *from tip to toe*, *head over heels*<sup>19</sup>. An intermediate type, comparable in meaning to *med top og tavl* is Da. *med hud og hår*, lit. ‘with skin and hair’.

For phonotactic reasons it is already clear that *\*-la-* in *\*tagla-* must be a suffix and cannot belong to the root. I will therefore tentatively reconstruct *\*dap-lo-* or *\*dop-lo-* as the PIE form behind PGmc. *\*tagla-* and OIr. *dúal*, with other derivatives *\*dap-* having reflexes in Germanic only. Relationship with Skt. *daśā-* ‘fringe’ of course remains a possibility, in which case the reconstruction is *\*dōk-lo-*<sup>20</sup>:

### 3.2 ODA. *SAKLÆ* ‘TO DROOL’, SW. DIAL. *SAKKEL* ‘DROOL’

A third relevant item is the Scandinavian verb for ‘to drool’, Da. *savle*, ODa. *saklæ*, Older Mod.Da. *sagle*, *sægle*, Sw. dial. *sakla*, *sagla*. It is probably denominal from Da. *savl* ‘drool’, Older Mod. Da. *sagel*, *sagle*, *sægle*, Sw. dial. *sakkel*. This noun has traditionally been connected to NWGmc. *\*sakkan-* ‘sink slowly, sag’ (> Da. *sakke* (*bagud*) ‘lag (behind), fall behind’), Mod.Icel. *sagga* ‘become moist’, Du. *zakken* ‘drop, sag’, and, outside Germanic, Slavic *\*sokŭ* ‘juice’. Semantically, though, it is as

<sup>18</sup> Nw. *topp*, Da. *top*, from PGmc. *\*tuppa-* (> ON *toppr* ‘tuft, lock of hair’, Far. *toppur* ‘crest’, OE *topp* ‘top’, OFris. *topp* ‘tuft’, OHG *zopf* ‘plait of hair’) can most easily be etymologized as a pseudo-etymological zero-grade of the root in *\*tappa-*, cf. the almost identical semantics of several forms and the lack of obvious alternative etymologies. The complexity of Germanic reflexes seem to confirm an origin in NW European IE and the “Germano-Celtic” vocabulary (Hyllested 2010).

<sup>19</sup> In lack of an established term for the latter type of idiom, Petr Kocharov (p.c.) suggests to introduce *polarindrome*.

<sup>20</sup> Kroonen (2013) as an alternative presents an inner-Germanic etymology according to which *\*tagla-* would be the diminutive to either *\*tahjan-* ‘to unravel’ or to *\*tōgan-* ~ *\*takkan-* ‘prickle, branch’.

close you can get to E Fris. *sabben* ‘to drool’, LG *sabbe* ‘drool; spit’, and further Du. *sabbelen* ‘to suck’, which contains PIE *\*sap-* (> PGmc. *\*saf/ppan-*, *\*safta-* ‘sap, juice ; moist’, Lat. *sapa* ‘must, new wine boiled thick’, Arm. *ham* ‘juice’ ; Kroonen 2013: 336). Here, too, I would therefore reconstruct an original labial and assert another derivative with *\*-lo-*, PIE *\*sap-lo-* > PGmc. *\*sakla-*, with no exact equivalents outside Germanic.

### 3.3 DA. *RAKLE* ‘CATKIN’, SW. DIAL. *RACKEL* ‘LONG THING ; TALL, SLIM PERSON’

Da. *rakle* ‘catkin, ament (flower cluster on trees)’ and Sw. dial. *rackel* ‘long thing; tall, slim person’ are etymologically obscure.<sup>21</sup> A PIE reconstruction *\*rop-lo-*, however, would render possible the establishment of a larger word-family, seeing that the Danish meaning comes close to Slavic *\*repjī* ‘burdock, arctium’ (> Ru. *repej* id., Ukr. *repyk* ‘sticklewort, Agrimonia’) and Alb. *rrap* ‘plane tree’ (whose fruits, *achenes*, are reminiscent of burdocks or round catkins of e.g. a hazel tree), while the Swedish meaning matches PGmc. *\*raftra-* ‘long, thin pole; rafter’ (> ON *raptr* ‘rafter’, OE *ræfter* ‘small beam’); both of these are derived from PIE *\*rep-* ‘stick to, pick up’ (> ON *ráfr* ‘roof on rafters’ ; Lat. *rapiō* ‘to snatch’, Gk. ἐρέπτομαι, Alb. *rjep* ‘tear off’, Lith. *ap-rėpiu*, *ap-rėpti* ‘to grasp’; LIV<sup>2</sup> 507), and formally they can even be equated with Lith. *rāplės*, *rėplės* f.pl., OPr. *raples* f.pl. ‘thongs’, albeit with an alternating ablaut grade. The original meaning of *\*rop-lo-*, *\*rep-lo-* would then be ‘snatcher’ and secondarily ‘burdock’, preserved best in Slavic, only later

<sup>21</sup> The unexpected *-k-* for *-g-* in Danish has not yet been explained. *Rakle* is not common in the singular; it is my impression that quite a few speakers only have the plural in active use. It is thus conceivable that *rakle* is a comparatively recent back-formation from the pl. *rakler* (Older Mod.Dan. pl. *rackle*), in turn from an older sing. *\*rakkēl* (a structure for which the modern pl. would also be *rakler*), corresponding to the Swedish form; cf. the parallel change of the sg. *skagel* into *skagle* below. In Danish, *-Vkke-* and *-Vgge-* are graphic renderings of the same pronunciation, so maybe *\*raggel* was simply a dialectal form of the type Zealandic *saggel* for *savl*, *sagl* (see the previous entry). If *\*raggel* was interpreted as the homophonous *rakkel*, a new plural *rakler* could arise. It is also possible that the irregular consonantism is due to contamination with dial./obs. Dan. *rakke*, also meaning ‘catkin, ament’, which is otherwise unrelated and constitutes a pair with the non-assimilated *ranke* ‘floral vine’. However, I find it more likely that the meaning of *rakke*, in turn, was influenced semantically by *rakkel*.

transferred to flower clusters of a burdock-like shape. A reconstruction *\*rap-lo-* would not match the root vocalism of Balto-Slavic, Greek or Albanian, so the Germanic *\*-a-* must reflect *o*-grade, and *e*-vocalism in the PIE root is ensured.

A problem compared to *\*hagla-*, *\*tagla-* og *\*sagla-* is of course that *\*rakla-* has PGmc. *\*-k-*, not *\*-g-*, an issue curiously repeating itself in Danish with *rakle* for expected †*ragle* (see the previous footnote). We will return to this variation below. Two further items point to a Germanic reflex *\*-akul-* from PIE *\*-apl-*:

### 3.4 PGMC. SKAKULA- ‘WHIPPLETREE; TOWING ROPE; SCHACKLE’

PGmc. *\*skakulaz* m. (> ON *skøkull* ‘car pole’, OE *sceacel*, *sceacul* ‘shackle, gyve, hobble’, MDu. *schakel* ‘chain link’; Da. *skagle* (< OMod.Da. *skagel*), Sw. *skakel*, Nw. *skokle* have two basic meanings ‘whippletree, double tree’ and ‘traces for harness horse, towing rope’; as a loanword in Fi. *kak(k)ula*, *ka(k)kuli* it only means ‘whippletree’. As stated in LÄGLOS (II : 18) the shape of the Finnish material is too unmarked to determine at what stage the borrowing took place; the limited geographical distribution—Finnish only, and mainly West Finnish dialects—tips the scale in favor of a late borrowing from Scandinavian/Norse rather than Proto-Germanic, maybe even as late as Old Swedish. Still, LÄGLOS leaves both possibilities open since theoretically *kak(k)ula* can also represent the Proto-Germanic accusative.

In either case, the meaning of the Finnish word is still important for our etymological purposes since it confirms that the meaning ‘whippletree’ is old in Scandinavian and apparently the only meaning exported eastwards. In combination with the many Danish meanings only referring to a piece of wood, both meanings must be reconstructed for common Scandinavian. The meaning ‘shackle’ is particularly West Germanic, but may have fallen within the Proto-Germanic range, too.

According to the standard etymology, PGmc. *\*skakulaz* is a diminutive formed with the suffix *\*-ula-* of the root in Nw.dial., Sw.dial. *skåk*, Sw. *skak* ‘towing rope; fettle for animals’ < *\*(s)kenk-*, cf. Lat. *cingere* ‘to gird’, Lith. *kinkýti* ‘to harness horses’<sup>22</sup>. Semantically it is an acceptable match, but there is no satisfying explanation as to why the nasal would

<sup>22</sup> Orel (2003: 332) instead derives *\*skakula-* from the root in the Sw. and ON verb *skaka* ‘shiver, quiver’ < PGmc. *\*skakana-*, with W Germanic meanings ‘run away’.

have disappeared in Germanic, and, again, it cannot be excluded that a contamination between two words has taken place.

Latin provides a clue. It is remarkable that Latin *capulus* displays a range of meanings similar to those of PGmc. *\*skakulaz*: a) ‘hackamore, headstall’, b) ‘halter for catching and fastening cattle; lasso’, c) ‘sword-hilt, handle’ and d) ‘coffin’. An original *\*skap-* can be found in Lat. *scāpus* ‘shaft’ and the otherwise disputed word for ‘shoulder (blade)’, *scapula*. *Scāpus* is already known to be related to PGmc. *\*skaftaz* ~ *\*skaftan-* m./n. ‘shaft; pole’ (> ON *skapt*, OE *sceaft*, OHG *schaft* id.), so the root *\*skap-* probably originally referred to a piece of wood connected to another part of a tool, rather than a rope-like device. I therefore suggest reconstructing PIE *\*skap-lo-* ‘shaft; double tree, whippetree’, which later, but already in NW IE, acquired the secondary meaning ‘traces for a harness horse; towing rope’ and ‘shackle(s)’. In Latin the missing *s-* in *capulus* can be explained by contamination with the near-synonymous *capistrum* ‘headstall, harness’, derived from *caput* ‘head’; thereby it is perhaps unnecessary to operate with *s mobile* in this case<sup>23</sup>. The meaning ‘coffin’ can be viewed as an enlargement of ‘handle’, but may simply be a historically different word, cf. the next item and especially the meanings of MLat. *cappellum* ‘chapel’, Lith. *kāpas* ‘grave’, Port. *campa* id.:

### 3.5 PGMC. *\*HAKULAZ* ‘CLOAK, MANTLE’

A very parallel formation can be seen in PGmc. *\*hakulaz* ~ *\*hakulōn* m./f. ‘cloak, mantle’ (> Goth. *hakuls*, OHG *hahhul*, ON *høkull*, OE *hacele* id., Da. *hagel* ‘cloak; (dial.) shawl’). It is usually regarded as a derivative of *\*hakōn* ‘hook, nook; buckle’ (Orel 2003: 154). I have earlier argued, however (Hyllested 2010), that MLat. *cappa*, synonymous to *\*hakulaz*, is not a medieval Semitic loanword as otherwise assumed, but inherited from PIE *\*kāpo-* which underwent the *littera*-rule whereby a long vowel plus a single consonant can be replaced by a short vowel followed by a geminate. Incidentally, the variant *cāpa* is attested in this

<sup>23</sup> *Capulus* itself can of course also be derived from *caput*, which, however, makes less sense when you consider the meanings b), c), and d). It should be noted that, regardless of etymological origins, Latin speakers probably had a feeling that both *capulus* in the meaning ‘headstall’ and *capistrum* belonged with *caput* synchronically, i.e. a contamination would not only be based on historical states of affairs.



case, and must represent the original form. *\*kāpo-* is not attested in the meaning ‘cloak’ outside Latin, but this problem can be solved by assuming that ‘cloak’ was not the primary meaning of the derivative. *\*kāpos* is already known as the PIE word for ‘piece of land; holy enclosure; garden’. It appears in Gk. κήπος ‘garden’, PGmc. *\*hōfa-* ‘(holy) enclosure’ (> ON *hof* ‘hill with holy place’) ~ *\*hōbō-* ‘piece of land’, and, with a PIE suffix *\*-i-sth<sub>2</sub>-* used for nomina loci<sup>24</sup>, in ORu. *kapište* ‘holy place, idol’ and Alb. *kophst* ‘garden’. This word-family was established by Witczak 2004; in Hyllested (2010) I added MLat. *cappellum* ‘chapel; holy enclosure’<sup>25</sup>; Lat. *castrum* in the meaning ‘fortification’, dim. *castellum*<sup>26</sup>; Lat. *campus* ‘field’<sup>27</sup>; *Capitōlium*, the name of various hills with holy places scattered around the Roman realms, most famously the one

<sup>24</sup> Cf. e.g. Gk. πλατάνιστος ‘plane-tree grove’, OHG *ewist* ‘sheepfold’, ON *vpzt* f. ‘fishing ground’.

<sup>25</sup> Said to be named after the small enclosure in the cathedral of Aachen where the relics of Saint Martin of Tours, including his cloak, are preserved. There can be no doubt that *cappellum* is formally secondary to *cappa*, but on the other hand there is no evidence that the meaning moved from a primary ‘cloak’ to a secondary ‘chapel’ in the middle ages (the French name of Aachen, *Aix-la-Chapelle*, is named after Charles the Great’s grave).

<sup>26</sup> Lat. *castrum* ‘fortification’ is most often understood as a result noun ‘piece cut out, strip of land’ corresponding to the homonym agent noun or instrument noun *castrum* ‘knife’ < PIE *\*kas-* ‘cut (out)’. Formally, though, *castrum* in the sense ‘fortification’ can equally well be derived from PIE *\*kap-i-sth<sub>2</sub>-ro-m* whose first two parts can be identified with the aforementioned nomen loci formation *\*kap-i-sth<sub>2</sub>-*. It would regularly yield Lat. *castru-* via syncope (of short vowel before *\*-st(r)-* as in *monstrum* ‘evil omen; monster’ < *\*monestrum*, *sēstertius* ‘two and a half; sesterce’ < *\*semi-s-tertios*; this happened in the 6<sup>th</sup>-5<sup>th</sup> c. BC, cf. Meiser 1998: 66) and subsequent assimilation (or loss, of *\*-T-* > Ø / *\_sT*, e.g. *Oscus* ‘Oscan’ < *Opacus* [Enn. Ann. 296], *asper* ‘raw’ < *\*ap-sper-*; cf. Meiser 1998: 117).

<sup>27</sup> Lat. *campus* ‘field’ can formally be identified with PIE *\*kámpos* ‘wave’ (> Gk. κάμπος), but semantically it corresponds more closely to the aforementioned *\*kāpos* which has otherwise left no trace in Latin. The most probable scenario is therefore that a contamination between *\*kampos* and *\*kāpos* took place on the way to Latin, resulting in a single word that retained the shape of the former and the meaning of the latter. There are signs that the contamination was not fully completed even in Vulgar Latin, since the original meaning of *\*kāpos* ‘enclosure’ seems to have been retained in Port. *campa* ‘grave’ (cf. also Lith. *kāpas* ‘grave’), although a homorganic nasal infix would be no less different that the one in the name of the region *Campānia* ~ Osc. *Kap(v)ans*, Etr. *capevane*, Gk. Καππανός; or It. *Campidoglio*, literally ‘the oil-plant fields’, but from Lat. *Capitōlium*.

in Rome itself<sup>28</sup>; the name of the Thracian fortification Καπιστούριον (cf. Duridanov 1987); the name of the temple *Secca di Capistello* for Diana and Aeolus from the 4<sup>th</sup> century BC on the Aeolic island of Lipari; *Capistrello*, an Italian place-name historically connected to Roman and Oscan fortifications; and the name of *Santiago de Compostela*, the famous catholic pilgrimage city in Galicia which also functioned as a holy place in Pre-Christian times<sup>29</sup>.

A double meaning ‘cloak, blanket, cover’ on one side and ‘piece of land ; (holy) enclosure’ on the other is typologically rather common, cf.:

- a) Skt. *namata-*, Av. *namāta-* ‘cloak’ next to PCelt. *\*nem-eto-* ‘holy enclosure’ (> OIr. *nemed*, Gaul. *Nemeto-*) and PGmc. *\*nem-ita-* (OLFr. *nimidas* and the Swedish farm-name *Nymden*); cf. also Latin *nemus* ‘holy enclosure’
- b) W *caen* f. ‘cover, blanket; skin, hide’ ~ Lat. *caulae* < *\*caholae* ‘enclosure’, PGmc. *\*hagan-* ‘garden’
- c) ON *kǫgurr* ‘blanket’ ~ *Kägra*, Swedish village- and farm name (Elmevik 1975)

In a parallel fashion, we can now establish:

- d) Mlat. *cappa*, *cāpa* ‘cloak, mantle’, PGmc. *\*hakula-* id. ~ PIE *\*kāp-o-*, *\*kap-ist-o-* ‘piece of land; (holy) enclosure’

Semantically this makes sense if the basic verbal meaning is ‘encircle, surround’ < ‘get hold of, hold’ < ‘catch, take’, cf. that the examples in a) above probably are derived from PIE *\*nem-* ‘to take’ (> PGmc. *\*neman-*

<sup>28</sup> *Capitōlium* can also go back to PIE *\*kap-i-sth<sub>2</sub>-*. According to Roman grammarians, an earlier form was *Capitōdium*, which renders possible an etymological segmentation *capi-tōd-*, cf. *cus-tōd-* < *\*cus-to-sd-* ‘he, who guards the treasure’; a derivative *\*kāp-ist-o-sd-jo-* would mean ‘he who belongs to the *\*kāp-ist-o-sd-*’, in turn ‘he who guards the *\*kāp-ist-o-*’, ‘the holy hill’, and a sequence *-sd-...-st-* could easily be subject of dissimilation, i.e. *\*kāpistosdiom* > *\*kāpitosdiom*. The guardian referred to can either be *Rēx Nemorēnsis* ‘the king of the holy grove’ (= Gaul. *Rigonemeti*), the goddess of the holy grove *Diana*, or (*Loucetios*) *Mars*.

<sup>29</sup> In Old Portuguese (the forefather of Modern Portuguese and Modern Galician), there was a confusion between the spelling variants <am> and <om> (cf. e.g. *orfom* ‘orphan’, found in king Duarte’s *Leal Conselheiro*, 1428-1438, from Romance ORFANU-) it is possible to assert an original *\*Campustella* or *\*Campistella*, simply meaning ‘the holy place’. Again, the variation between forms with nasal and forms without point to confusion between *\*kampos* and *\*kāpos* at a late stage.

). Correspondingly, d) can originally be identical to the verbal root *\*kap-* ‘to catch’ nown from PGmc. *\*habēna-* ‘to have’ and Lat. *capīre*.

I conclude that PGmc. *\*hakula-* ‘cloak’ can be derived from PIE *\*kap-lo-*, a derivative of the same root as in MLat. *cappa*, *cāpa* ‘cloak’.

#### 4 Conclusions

We end up with six potential examples of the postulated development PIE *\*-apl-*, *\*-opl-* > PGmc. *\*-agl-*, *\*-akl-*, *\*-akul-*—and, crucially, no obvious counterexamples:

- PGmc. *\*hagla-* ‘hail’ < PIE *\*kōp-lo-* (~ Late PIE *\*kōp-ro-* ‘crops, grain’ > PGmc. *\*habra-* ~ *\*hagra-* ‘oats’)
- PGmc. *\*tagla-* ‘(horse)tail’ < PIE *\*dōp-lo-*, *\*dap-lo-* (~ Da. *tave*, Sw. *tafse* ‘tuft’, PGmc. *\*tappan-* ‘tap’)
- PGmc. *\*sagla-* ‘drool (sb.)’ < PIE *\*sap-lo-* (~ E Fris. *sabben* ‘to drool’, LG *sabbe* ‘drool; spit’, PGmc. *\*saf/ppan-*, *\*safta-* ‘sap, juice; moist’; Lat. *sapa* ‘must’)
- PGmc. *\*rakla-* ‘catkin’ < PIE *\*rop-lo-* (~ PGmc. *\*raftra-* ‘rafter’, Lith. *rāplės* f.pl. ‘thongs’, Ru. *repej* ‘burdock’, Alb. *rrap* ‘plane tree’)
- PGmc. *\*skakula-* ‘whippletree; towing rope; schackle’ < PIE *\*skap-lo-* (~ PGmc. *\*skafta-* ‘shaft; pole’, Lat. *capulus* ‘hackamore; lasso; handle’, *scapula* ‘shoulder’, *scāpus* ‘shaft’)
- PGmc. *\*hakula-* ‘cloak, mantle’ < PIE *\*kap-lo-* id. (~ MLat. *cappa*, *cāpa* ‘cloak, mantle’)

The extra *-u-* in the two final examples has been explained by Kümmel (2004) as a reflex of partly analogical, partly regular developments within the paradigms of Germanic nouns with a stop followed by a sonant in the stem.

The division into three items with voiced spirants and three other items with unvoiced stops also needs an explanation. Such an alternation is reminiscent both of Verner’s Law (voiced vs. unvoiced) and parts of Kluges’ Law (retention of an original unvoiced stop before sonant; here *-l-* which is retained, and not a nasal assimilated into the stop) and could reflect original accent alternation in exactly this position. The development itself can perhaps be characterized as a sort of assimilation, if *\*-l-* at the point in question was of the thicker, velar kind.

Future studies can perhaps reveal whether other kinds of delabialization processes in Germanic (see e.g. Markey 1979) can be explained as the result of similar processes. It is important to note that the situation outlined in this article does not necessarily contradict Kortlandt’s (1997: 48) hypothesis that labiovelars went in the opposite direction and became real labials next to sonorants.

## References

- Beekes, Robert S.P., 2010: *Etymological Dictionary of Greek*. Leiden: Brill.
- Bjorvand, Harald & Fredrik Otto Lindeman (eds.), 2001: *Våre Arveord. Etymologisk Ordbok*. Oslo: Novus.
- Brøndum-Nielsen, Johannes, 1935: *Gammeldansk Grammatik i Sproghistorisk Fremstilling*. Copenhagen: J.H. Schultz.
- Duridanov, Ivan, 1987: “Die geographische Terminologie indogermanischer Herkunft im Thrakischen und Dakischen”. – Wolfgang Meid (ed.): *Studien zum indogermanischen Wortschatz*. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft. Pp. 29-34.
- Elmevik, Lennart, 1975: “Køgurr och Kägra”. – *Namn och Bygd* 63: 38-49.
- Griffiths, Alan, 2006: “Rune-names: The Irish Connexion”. – In M. Stoklund, B. Holmberg, M. Lerche Nielsen and G. Fellows-Jensen (eds.): *Runes and their secrets*. Copenhagen: Museum Tusculanum Press. Pp. 93-101.
- Hellquist, Elof, 1898: “Om nordiska verb på suffixalt *-k*, *-l*, *-r*, *-s*, *-t* samt av dem bildade nomina”. – *Arkiv för Nordisk Filologi* n.f. 10: 1-46 and 136-194.
- Hofstra, Tette, 1995: “Beziehungen zwischen Germanen und Ostseefinnen vor der Hansezeit”. – *Tijdschrift voor Skandinavistiek* 16, no. 2: 85-103.
- Hyllested, Adam, 2010: “Saint Martin of Tours, the Capitolin Triad, and Latin Reflexes of PIE \**kāp-o-*, \**kap-ist-o-* ‘piece of land; (holy) enclosure’ “. – Foredrag afholdt på *Greek and Latin in an Indo-European Perspective III* i Bratislava, Slovakiet, juli 2010.
- Hyllested, Adam, 2012: “*Hagl*, *tavl* og *skagle*. Flere undtagelser til den germanske lydforskydning? – in Hans Frede Nielsen (ed.): *Ældre germansk sprogshistorie* [= *Mindre Skrifter Blaa Serie* 29]. Odense: University Press of Southern Denmark. Pp. 5-12.
- Hyllested, Adam & Paul S. Cohen, 2007: “Monophthong for expected v-diphthong in Greek”. – Matthew McCullagh, Coulter George, Benedicte Nielsen, Antonia Ruppel & Olga Tribulato (eds.): *Greek and Latin from an Indo-European Perspective* [= *Proceedings of the Cambridge Philological Society Supplementary Volume* 32]. Cambridge: Cambridge Philological Society. Pp. 9-18.
- Kerkhof, Peter, 2012: *Suffix variation in the PGmc. l-suffixes and the ablaut of the PIE l-stems*. PhD dissertation from Leiden University.
- Kortlandt, Frederik, 1997: “Labials, Velars and Labiovelars in Germanic”. – *NOWELE* 30:45-50.
- Kroonen, Guus, 2013: *An Etymological Dictionary of Proto-Germanic*. Leiden: Brill.

- Kümmel, Martin Joachim, 2004: "Ungeklärtes *u* neben Liquida in Germanischen Nomina". – A. Hyllested, A.R. Jørgensen, J.L. Larsson og T. Olander (eds.): *Per aspera ad asteriscos. Studia Indogermanica in honorem Jens Elmegård Rasmussen anno MMIV*. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft. Pp. 291-303.
- Lockwood, W.B., 1961: *The Faroese Bird Names*. Copenhagen: Munksgaard.
- LÄGLOS = Andries D. Kylstra, Sirkka-Liisa Hahmo, Tette Hofstra & Osmo Nikkilä: *Lexikon der älteren germanischen Lehnwörtern in den indogermanischen Sprachen I-III*. 1991-2012.
- LIV<sup>2</sup> = Helmut Rix (ed.): *Lexikon der Indogermanischen Verben*. 2nd ed. 2001. Wiesbaden: Reichert.
- McManus, Damian, 1988: "Irish letter-names and their Kennings". – *Ériu* 39: 127-168.
- Markey, T.L., 1979: "Delabialization in Germanic". – *Journal of Indo-European Studies* 7: 199-211.
- Meiser, Gerhard, 1998: *Historische Laut- und Formenlehre der lateinischen Sprache*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Olson, Johan Emil, 1904: *Östgötalagens Ljudlära*. Lund: Gleerupska Universitetsbokhandeln.
- Orel, Vladimir, 2003: *A Handbook of Germanic Etymology*. Leiden / Boston: Brill.
- Rasmussen, Jens Elmegård, 1984 [1999]: *Miscellaneous Problems in Indo-European Languages III*. – *APILKU* 4: 135-149 [reprinted in J.E. Rasmussen: *Selected Papers on Indo-European Linguistics*, 1999. Museum Tusculanum Press. Pp. 146-155].
- Rydquist, Johan Erik: 1868: *Svenska Språkets Lagar I-VI*. Stockholm: Norstedt.
- Seip, Didrik Arup, 1931: *Norsk Språkhistorie till omkring 1370*. Oslo: Aschehou & co. (W. Nygaard).
- Tamm, Fredrik August, & Adolf Noreen, 1890-1905: *Svensk Etymologisk Ordbok* 1-8. Akademiska boktryckeriet E. Berling.
- Wessén, Elias, 1965: *Svensk Språkhistoria I-III*. Stockholm / Gothenburg / Uppsala: Almqvist & Wiksell.
- Wiklund, K.B., 1917-1920: "Die ältesten germanischen Lehnwörter im finnischen. – *Indogermanische Forschungen* 38: 48-114. (61-62)
- Witczak, Krzysztof Tomasz, 2003: *Indoeuropejskie nazwy zbóż*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
- Witczak, Krzysztof Tomasz, 2004: "Albanian *kophst* 'garden' and its cognates". – *Studia Etymologica Cracoviensia* 9: 179-787.

## Two Issues on Indo-European Substrates in Slavic<sup>1</sup>

### Abstract

Georg Holzer's (1989) identification of an Indo-European substrate language in Slavic – *Temematian* – is correct, as also argued by Kortlandt (2004). However, Kortlandt's skepticism regarding the specific sound development PIE \**ṛ* > Temematian \**Ro* has proved unfounded. Holzer did not give any examples of substantivized past passive participles, an otherwise very common word-type which, if containing a liquid in the root, would surface as \**CRod̥* or \**CRon̥* in Temematic words in Slavic. It is suggested that at least LCS \**plod̥* 'fruit' and \**grozd̥* ~ \**grozn̥* 'grape; cluster', perhaps also \**drozd̥* 'thrust', are Temematic borrowings. Another possible Temematism is *zabar* 'oats' in the Alanic (Iassic) word-list.

Trubachëv's hypothesis that Baltic and Slavic \*-*st*- can reflect PIE \*-*ḱ*- in lexems from another Indo-European substrate is rejected, seeing as in all of his examples, *st* can be accounted for in other, more transparent ways.

### 1 The problem of identifying related substrate languages

Virtually all specialists would agree that Indo-European languages are newcomers in Europe which replaced most of the indigenous languages in several steps from the Bronze Age onwards. To which extent this lin-

---

<sup>1</sup> This article forms part of a larger paper, "Indo-European Substrates in Slavic Revisited" that was presented at the 17<sup>th</sup> Conference of Scandinavian Slavists in Copenhagen, August 2007; at the 19<sup>th</sup> Annual UCLA Indo-European Conference, 2 November, 2007; and at the 14th International Congress of Slavists in Ohrid, Macedonia, 11 September 2008. Most of the points in section 3 were printed in Hyllested 2004, while the etymology for *plod̥* only was presented at the XVI Conference of Scandinavian Slavists in Uppsala, August 2004.

guistic replacement reflects actual human migrations, let alone entire “population shifts”, is another (albeit related) matter and a matter of debate – but the great changes that took place may in any case be characterized as “Indo-Europeanization”. The encounter between indigenous Palaeo-European cultures and what we now define as Indo-European culture, however, was most certainly not the only interaction that took place. By the time of arrival in Europe, the Indo-European languages and populations were already fragmented and scattered in many directions – some were still immediate neighbours and must have exchanged lexemes to some extent.

Convergence between closely related languages in the past constitutes one of the greatest challenges for historical linguists, especially when one of them is neither directly attested nor has any attested descendants. The bulk of the vocabulary is simply either directly inherited or borrowed from closely related dialects who got their terms ultimately from the same precursor. The consequence is typically that linguists overlook the language which disappeared because loanwords from it are not interpreted as preserved traces of a forgotten language, but merely irregular varieties of inherited material in the related language that borrowed them.

## 2 Holzer’s *Temematisch*

Holzer (1989) suggested an extinct and otherwise unknown Indo-European language, which he calls “Temematisch” (Eng. *Temematian*) to account for a variety of irregular consonant correspondences in Slavic. He ascribes 62 lexemes in Slavic to this hypothetic extinct language – basically all lexemes with troublesome or unsatisfying etymologies. The language name is an acronym based on the asserted developments: PIE *tenues* (unvoiced stops) became *mediae* (voiced stops), and *mediae aspiratae* became *tenues*. Other Temematian features, according to Holzer, are that zero-grades of liquids surface as *\*-Ro-*, and that long vowels become shortened before sonant.

PIE	Temematic (in Slavic)	Normal Balto-Slavic corresp.
<i>*b<sup>h</sup>, *d<sup>h</sup>, *g<sup>h</sup></i>	<i>&gt; *p, *t, *k</i>	<i>~ *b, *d, *g</i>
<i>*p, *t, *k</i>	<i>&gt; *b, *d, *g</i>	<i>~ *p, *t, *k</i>
<i>*r, *l</i>	<i>&gt; *ro, *lo</i>	<i>~ mostly *il, *ir</i>
<i>Ṽ/_R</i>	<i>&gt; V</i>	<i>~ distinct reflexes</i>

Among Holzer's best examples (in my opinion) are:

PIE *\*suo-poti-* 'one's own lord' (cf. Vedic *svápati-* id.) > LCS *\*svobodī* 'free' (for expected †*svopotī*)

PIE *\*d<sup>h</sup>el-(ent-)* 'suck, nurse' (cf. Latvian *dīle* 'sucking calf') > LCS *\*telē* 'calf' (for expected †*delē*)

PIE *\*b<sup>h</sup>rs-o-* 'grain, crops' (cf. CSL. *brašino* 'flour', ON *barr* 'barley') > LCS *\*proso* 'millet' (for expected †*\*brīsū* or †*\*brīso*). Compare now Toch. B *proksa* (pl.) 'grain', but cf. also LCS *\*būrū* 'millet' < *\*puH-ro-*

Kortlandt (2004) evaluated Holzer's theory positively, also in a historical context, but had one minor remark: He did not consider Holzer's specific Temematan development *\*R̥ > Ro* more than theoretically possible, pointing to the fact that the evidence is limited to five lexemes, and that in none of these zero-grade is to be expected:

- 2. LCS (via Tem.) *\*proso* < PIE *\*b<sup>h</sup>rs-o-* ~ Lat. *far*
- 6. LCS (via Tem.) *\*loboda* < PIE *\*l̥-podā* ~ Lat. *olor*
- 10. LCS (via Tem.) *\*krotū* < PIE *\*g<sup>h</sup>rd<sup>h</sup>o-* ~ Goth. *garda*
- 27. LCS (via Tem.) *\*prokū* < PIE *\*b<sup>h</sup>rg<sup>h</sup>o-* ~ Goth. *baírgan*
- 30. LCS (via Tem.) *\*slobodū* < PIE *\*s̥l̥-poti-* ~ ON *salr*

While I agree with Kortlandt that this is one of the weaker Temematan correspondences judging from Holzer's own material, I find it remarkable that no past passive participles in *\*-tó-* are represented since they are common and constitute the zero-grade environment *par excellence*. According to Holzer's rules, such participles if borrowed from Temematan would end in LCS *\*-dū*. If the stem ended in a liquid (perhaps followed by a laryngeal) it would be easy to check if Holzer's rule is correct since in that case they would have the shape *\*CRodū* corresponding to regular Slavic *\*CbRtō* and PIE *\*CR̥(H)-tó-*. Since the participles formed part of the conjugational system and could be formed productively as adjectives and substantives even in the daughter-languages, we would expect at least a couple of those containing liquids in the root to show up in the Temematan material – but of course they can only be identified if Holzer was right that their zero-grade was markedly different from the regular Slavic one.



2.1 LCS \**PLODŮ* 'FRUIT'

The Slavic word for 'fruit', OCS *plodŭ*, although of disputed origin (see e.g. Bezlaĵ 1995), is most often taken for a loanword from PGmc. \**blaða-* 'leaf, ? fruit' < PIE ppp. \**bʰl̥a₁-tó-*, ultimately from the root \**bʰel-* 'swell'. However, the exact phonological representation remains a problem: Why is PGmc. \**b-* replaced by *p-*, and why is the vowel -*o-*?

Shevelov (367) equated it with Celtic forms like Ir. *loth* 'foal' and Welsh *llwdu* 'young person', while Bezlaĵ (1995) proposes a back-formation from PIE \**pled-men* (with voice assimilation) < \**plet-men* 'rope; thread', whence the meaning 'lineage; offspring' (LCS \**plemē*); he refers to typological parallels with double meanings such as Skt. *tantu* 'thread; offspring'; Slov. *pasma* 'race; train column'; and SCr. *lòza* 'rope; fruit tree'.

Curiously, a PIE participle \**bʰl̥(H)-tó-* of this root, where a zero-grade is obviously expected, would in fact yield Temematic \**plodŭ*. It corresponds formally to PGmc. \**bulð-i/ō-* which is attested e.g. as Dan. *byld* 'abscess'. The root could also have contained a laryngeal because it is the loss of long-short distinction is another Temematic feature ( \**bʰl̥h₁-t'*: -*R̥H-* > -*Ro-*; Kortlandt 2004).

The meaning 'fruit' in PGmc. occurs only with full-grade and feminine gender: OE *blæd* f., MLG *blāt* f., but on the other hand we encounter a similar semantics in zero-grades of other extended root-forms like Sw. dial. *böljon* 'blueberry' (< PGmc. \**bulyi-* < \**bʰl̥-ġʰ-í-*), Dutch *bolster* 'fruit shell, husk', and Gk. φλυδαρός 'overripe'.

2.2 LCS \**GROZDŮ* 'GRAPE; ? CLUSTER'

The same semantic sphere – fruticulture and viticulture – remarkably features another good candidate for a substantivized Temematic participle. LCS \**grozdŭ* likewise has an unexpected -*Ro-* while ending in -*dŭ*. The traditional etymology, deriving it from PIE \**ġʰerzd-* 'barley', does not explain the synonym *groznŭ*. However, nasal participles are preserved in Slavic, so the variation \**grozdŭ* ~ \**groznŭ* is yet another sign that we have encountered a Temematic participle. They can be completely identical to PGmc. \**hursta-* 'shrubbery' (< \**kʷrs-tó-*) and OIr *crann* (< \**kʷrs-nó-*) 'tree' (meaning originally 'vine'?) respectively. But there is also the possibility that the Russian meaning 'cluster' represents an archaism, and that 'grape' is secondary – in that case, the word could

be identical to Latv. *gūrste* ‘linen bundle’ and OCS *gr̥stb* ‘handful’ < PIE *\*gr̥t-sti-*.

### 2.3 LCS *\*DROZD̥b* ‘THRUST’

A Slavic *\*drozd̥b* (e.g. Ru. *drozd*) ‘thrust’ exists alongside the more widespread variant *\*trozd̥b* (OCS *trozdŭ*). However, this case is less certain because forms from other Indo-European languages show that a) *\*-d̥b* is not participial, but comes from an original PIE sequence *\*-d-o-*, and b) while zero-grade forms do occur, like Lat. *turdus* (< *\*t̥r̥zdos*), the more closely related Lith. *strāzdas* with o-grade (< *\*strozdo-*) shows that the sequence *\*-rV-* is original in the full-grade, and Slavic *\*-ro-* does therefore not have to reflect a Temematian zero-grade. However, *\*drozd̥b* would be the expected Temematian outcome of the word in any case.

### 2.4 ALANIC (IASSIC) *ZABAR* ‘OATS’

Johnny Cheung (p.c.) has suggested that *zabar* in the Alanic (Iassic) word-list is simply a rendering of Hungarian *zab* which is loanword from Slavic. This solution, however, does not satisfactorily account for the second part of the word *-ar* – although Cheung suggests that it might be Hung. *ár* ‘price’. Incidentally, *zabar* is the only word on the list without a counterpart in Modern Ossetic. However, all other words in the left column are consistently Iranian, so it seems reasonable to interpret *zabar* ‘oats’ as the corresponding Iassic word for ‘oats’.

Holzer includes Slavic *\*zob̥b* in his list of Temematian words. The Iassic word could either be inherited from Proto-Iranian or reflect yet another Temematian agricultural term in Slavic (subsequently borrowed into Iassic). In either case, *zabar* would be semantically and morphologically identical to the West Indo-European ‘oats’-words (PGmc. *\*habran-*, PCelt. *\*korkio-*) versus the more archaic formations and meanings in the East (Hitt. *kappar* ‘vegetables’, Skt. *śāpa-* ‘drifting reed’), thus reflecting a morphological and semantic innovation pointing rather clearly to an agricultural specialization that follows the earliest dissolution of the Indo-European dialects and migrations into Europe. Iassic).

### 3 Trubachëv's substratum reflex *-st-* < *\*-ĥ-*

Another Indo-European substratum language, more specifically an extinct Balto-Slavic dialect or Satem dialect, was established by Trubachëv (e.g. 1991) mostly to account for the numerous discrepant dorsals, i.e. kentum-like reflexes of PIE palatal stops. However, another reflex ascribed to ancient dialect convergence is Baltic or Slavic *st* for expected *s* from PIE *\*ĥ*. Three examples are mentioned by Andersen (2003):

#### 3.1 LITH. *STIRNA*, LATV. *STIRNA* 'DEER'

Lith. *stirna*, Latv. *stīrna* 'deer' are seen as substratum reflexes of PIE *\*ĥer-n-*, cf. PSl. *\*sirnā*. However, not only Slavic, but also Old Latvian has *s-*, cf. pl. <*Ssirnos*>. The forms with *st-* thus appear recent and are likely to have been influenced by German in medieval times, cf. OHG *stiora* (NHG *Stier*) 'bull' or OHG *stirna* 'forehead; skull pad on war-horse' (< 'blaze, characteristic forehead')<sup>2</sup>.

#### 3.2 LITH. *TŪKSTANTIS*, LATV. *TŪKSTOTIS* 'THOUSAND'

Lith. *tūkstantis*, Latv. *tūkstotis* 'deer' is another of Trubachëv's prominent examples, but here *-st-* stands in medial position and reflects the regular Baltic development of the Balto-Slavic cluster *\*-s-ts-* with *\*-ts-* as the intermediate stage between *ĥ* and its final unmarked outcome *s*. Pedersen (1942) thought that this only happened in position before front vowel whereas the regular outcome in front of back vowel is *-sk-*. However, note Lith. *laiškas* 'leaf' vs. OCS *listŭ* id. This Balto-Slavic word is a more obvious source for Fi. *lehti* 'leaf' (< *\*lešte*) and Mari *lištaš*, *laštaš* 'id.' than (NW) PIE *\*b<sub>l</sub>h<sub>1</sub>-to-/\*bleh<sub>1</sub>-to-* as suggested by Koi-vulehto (1995).

<sup>2</sup> Irregular initial position seems to be common in names for big mammals. Even Lith. *stumbras* 'bison' has an unexpected *st-* although from PIE *\*ĝ-* Latv. *sumbrs* ~ CSL. *žqbrŭ*; cf. parallels like OIr. *fearb* ~ *earb* 'deer; cow'; ON *þjórr* ~ *stjórr* 'bull' (PGmc. *\*þeura-* ~ *\*steura-*).

3.3 OLD PRUSSIAN *PARSTIAN* 'PIG'

The sequence in OPr. *Parstian*, rendered <prastian> (Elbinger Vocabulary 686) is to be segmented *pars-tia-n* where *-tia-(n)* forms a diminutive, as in e.g. *werstian* 'little bull' calf (EV 674) vs. Lith. *veršis* 'calf', *wosistian* <wolistian> 'kid' (EV 677) vs. Lith. *ožys* 'goat', and *eristian* 'lamb' (E 681) vs. Lith. *ėras* id. Thus, only *-s-* and not the entire sequence *-st-* reflects *\*-k̑-* in (NW) PIE *\*pork̑o-*.

According to Leskien (1891: 583) and Skardžius (1943: 332) the suffix shape is *-istian* and may be connected with Lith. *-iščia-* (see also Schmalstieg 2003: 271), assuming that forms without *-is-* have undergone dissimilation, e.g. *werstian* < *\*versistjan*. It remains that *-t-* is suffixal (whether it belongs with *-s-* or not) and cannot reflect PIE *\*-k̑-*.

## References

- Andersen, Henning, 2003: "Slavic and the Indo-European Migrations". – Henning Andersen (ed.): *Language Contacts in Prehistory*. Amsterdam/Philadelphia: John Benjamins. Pp 45-76.
- Bezljaj, France, 1995: *Etimološki Slovar Slovenskega Jezika* III: P-S. Re-edited by M. Snoj & M. Furlan. Ljubljana.
- Holzer, Georg, 1989: *Entlehnungen aus einer bisher unbekannten indogermanischen Sprache im Urslavischen und Urbaltischen*. Vienna.
- Hyllested, Adam, 2004: Review of Andersen 2003. – *Acta Linguistica Hafniensia* 36: 179-184.
- Koivulehto, Jorma, 1995: "Zur indogermanisch-germanischen Kontinuität in der Nachbarschaft der Finnougrier". – *Der Ginkgo-Baum. Germanistisches Jahrbuch für Nordeuropa* 13. Helsinki. Pp. 116-137.
- Kortlandt, Frederik, 2004: "An Indo-European Substrate in Slavic?". – Alfred Bammesberger & Theo Vennemann (eds.): *Languages in Prehistoric Europe*. Heidelberg. Pp.
- Leskien, August, 1891: *Die Bildung der Nomina im Litauischen*. Leipzig: Königl. sächs. Gesellschaft der Wissenschaften.
- Pedersen, Holger, 1942: "Et baltoslavisk problem". – Rosally Brøndal, Viggo Brøndal, Christen Møller & Hedvig Olsen: *In Memoriam Kr. Sandfeld*. Copenhagen: Gyldendalske Boghandel Nordisk Forlag. Pp. 184-194.
- Schmalstieg, William R., 2003: "An isogloss uniting Baltic, Slavic, Germanic". – Alfred Bammesberger & Theo Vennemann (eds.): *Languages in Prehistoric Europe*. Heidelberg. Pp. 261-278.
- Shevelov, George Y., 1964: *A Prehistory of Slavic*. Heidelberg: Carl Winter.
- Skardžius, Pranas, 1943: *Lietuvių kalbos žodžių daryba*. Vilnius: Lietuvos mokslų akademija.
- Trubachëv, Oleg, 1991: *Etnogenez i kul'tura drevneisikh slavjan*. Moscow: Nauka.



# The Story of ‘time’: The Etymology of Finnish *aika*

(with an excursus on *aita* ‘fence’ and Balto-Slavic *v*-prothesis)

## Abstract

The traditional etymology deriving *aika* from Germanic *\*aiwa-* ‘time’ must be rejected because PGmc. *\*-w-* would not be rendered as Balto-Fennic *\*-k-*. Instead, it is a borrowing from a different Germanic derivative from the same PIE stem, namely *\*ajuka-* ‘life, age’, an old thematic noun forming the basis of the adjective *\*ajuki-* ‘eternal’, which can be reconstructed on the basis of Old English *ēce* and Gothic *ajuk-* in *ajukdūps* ‘eternity’. The development of *\*-u-* > *\*-e-* in Middle Proto-Fennic and subsequent loss in Late Proto-Fennic is regular in the medial position of such loans.

While BF *\*aita* ‘fence’ can be reconstructed for Fenno-Ugric, it might ultimately be a late Indo-European borrowing from *\*h<sub>1</sub>oi<sub>2</sub>to-* ‘delimitation, demarkation’ that yielded the Celto-Germanic word for ‘oath’, Lith. *vietà* ‘place’ and Slavic *věťb* ‘council; convention; oath’. BF *\*aita* and *věťb* then form a pair systematically corresponding to *\*aika* vs. Slavic *věkō* ‘time’, indicating a development of initial PIE *\*(H)oi-* to Balto-Slavic *\*vai-* in sandhi contexts.

## 1 Finnish *aika* ‘time’ cannot come from PGmc. *\*aiwa-*

Kari Liukkonen (1999: 18-19) justifiably characterizes the traditional etymology deriving *aika* from Germanic *\*aiwa-* ‘time’ as phonetically impossible: Germanic *-w-* would not be rendered as Balto-Fennic *\*-k-*. Instead he proposes that Balto-Fennic *\*aika* ‘time’ (> Fi. *aika*) is a rendering of Baltic *\*eigā* ‘course (of events)’ (> Lith. *eigà*), an old derivative of the verb *eĩti* ‘to go’, cf. also *eigoje* ‘during’ (< PIE *\*h<sub>1</sub>ei-* ‘go’).

However, this does not work either. Proto-Baltic *\*ei* is regularly retained as Balto-Fennic *\*ei* which only later develops into *ai* in South Estonian and Livonian. Larsson (forthc.) states:

“It must first and foremost be clarified that PBalt. *\*ai* can indeed yield EastBaltic *ie* (e.g. Lith. *dieveris* ‘brother-in-law’ and ORu. *děverø*, Gk. *dāēr*, Lat. *laevir*, Arm. *taygr* [...]) Another key example is Fi. *taivas* ‘heaven, sky’ which is generally said to be a borrowing from Balt. *\*deiuas* (Lith. *diēvas*, Latv. *dievs*, OPr. EV *deiwis* ‘god’) [...] However, this example is better explained [...] as] an early loan from Indo-Iranian, i.e. IIr. *\*daiuas*”.

One could add that the study of semantic fields speaks for an Indo-Iranian origin. Fi. *jumala* ‘God’, attested already in Ottar’s account as ON *Jómáli* is a derivative from an old name for the sky, formed with the productive nomen loci suffix *\*-la*. The base word *juma* is of Indo-Iranian origin and identical to Skt.*dyúman-* ‘sky’ (derived from PIE *\*deǵ-*, *\*dǵeu-*, famous for its occurrence in names of Indo-European gods. Together with *taivas*, religion already seems to constitute a visible semantic field among Indo-Iranian loans.

A late development of initial *\*e-* to *\*a-* (not only when it forms part of diphthongs) is known from Lithuanian (cf. dial. *aĩgis*), but not from Baltic as such, and since cognates of *aika* are found all over Balto-Fennic, the word must be a very old borrowing. One might of course conjecture that *\*ei-* regularly yielded *ai-* exactly in initial position, but there are not many examples (Andersen 1996 does not mention any), so such an argumentation would be circular. We have a motivation for initiating a search for alternative etymologies.

## 2 The oath and the fence – an origin in Balto-Slavic?

One possibility is that *aika* could be older than Baltic, going back to Balto-Slavic. Fi. *aita* ‘hedge, fence’, with attested cognates in all Balto-Fennic languages but Livonian, is of a similar structure and relevant in this context. Koivulehto (1973) suggested that *aita* reflects Proto-Germanic *\*aiþa-* ‘oath’, mentioning the semantic parallel in Greek ἔρκος ‘fence, hedge’ next to ὄρκος ‘oath’. Unfortunately he withdrew the etymology in his 1999 version, but the idea was not bad. The common semantic denominator would have been something like ‘delimitation’, ‘demarkation’ or ‘restriction’.

Germanic *\*aiþa-* is otherwise regarded as a Celto-Germanicism (Hyllested 2010), corresponding to Celtic *\*oito-* (> Old Irish *óeth*) and usually interpreted as a lexicalization of PIE ‘a walk’ (< *\*H<sub>1</sub>oi-to-*) because of the ON expression *ganga eið*, but this parallel does not really hold water since the meaning ‘walk’ is lost when one removes *ganga*. Another possibility is that we are dealing with PIE *\*h<sub>3</sub>eit-* ‘fetch’, established by Tichy (2004) on the basis of Gk. *οἶσσομαι* ‘to fetch, take along’ and Lat. *ūtor* ‘to use’ and supported by Melchert (2007) who supplies Cuneiform Luvian *hizza(i)-* ‘fetch’ (in collocations with a motion verb). As a semantic parallel we could recall PIE *\*kag<sup>h</sup>-* ‘to hold’ > Gmc. *\*hagan-* ‘garden’.

BF	Gmc.	Baltic	Slavic	IE
<i>*aika</i>	<i>*ajuka-</i> <i>*ai-h-ta-</i>	<i>*vaika-</i>	<i>*věkǫ</i>	<i>*H<sub>2</sub>oiu-go-</i> , <i>*H<sub>2</sub>oiu-ko-</i>
<i>*aita</i>	<i>*aiþa-</i>	<i>*vaita-</i>	<i>*větǫ</i>	<i>*H<sub>1</sub>oi-to-</i> or <i>*H<sub>3</sub>eit-o-</i>

We already know that PIE *\*(H)oi-* can yield Baltic *\*vai-*, e.g. *\*oiH-no-s* ‘1’ > Lith. *vienas* and *\*Hoǵstro-* > Lith. *aistra-* ‘vehement passion’ ~ Liv. *aistar* ~ dial. *vistar* ‘pimple’, a Baltic loan (cf. both meanings of the Greek cognate elsewhere in this publication). What is the Slavic reflex of such an initial diphthong? There seems to be only a single relevant, but contradicting, example: *jědro* ‘disease mark on tree’ ~ in Latv. *idra*, *idrs* which happens to be related to the root of the aforementioned Baltic word < PIE *\*Hoid-ro-* ~ *\*Hoǵd-tro-*. PIE *\*oiH-no-* ‘1’ is only attested in the zero-grade as *jed-inǝ* ‘1’ and *ino-roǵǝ* ‘unicorn’.

Finnish *aika* and *aita* form a curious double pair with Slavic *\*věkǫ* ‘time’ and *\*větǫ* ‘council; convention; oath’ (cf. also Lith. *vieta* ‘place’). I see two possibilities:

- 1) Slavic *\*vě-* could be the regular reflex of some subset of PIE *\*(H)oi-*, e.g. *\*H<sub>3</sub>oi-* (*\*H<sub>3</sub>* being a labial consonant)? This solution, however, would be based on circular argumentation since none of the etymologies are properly established. More examples are needed.
- 2) Slavic *\*vě-* < could have arisen in sandhi of PIE *\*(H)oi-* in cases when a rounded element proceeded – since a rounded element follows as well. Cf. *so-větǫ* ‘convention’ (Ru. *soviet*). One weakness of this possibility is that *jědro* is not an obvious 2<sup>nd</sup> member of compounds.



It is important to note that Baltic *\*vaitā* 'place' and *\*vaika-* 'child' can come from PIE *\*(H)oi-* without a problem. Could Balto-Slavic *\*(H)oi-* yield Slavic *\*vě-* in sandhi for expected *\*jě-*, parallel to regular Baltic *\*vai-*? It is of course theoretically possible that the Slavic words are borrowings from Baltic, but the conclusion would be the same.

### 3 Rather ← PGmc. *\*ajuka-* 'time' → *\*ajuki-* 'eternal'

As mentioned above, Liukkonen justifiably characterizes the traditional etymology *aika* < Germanic *\*aiwa-* 'time' as phonetically impossible: *\*w-* would not be rendered as Balto-Fennic *\*-k-*. However, it has been overlooked that Germanic possesses a related derivative, PGmc. *\*ajuki-*, albeit an adjective meaning 'eternal', which can be reconstructed on the basis of Old English *ēce* and Gothic *ajuk-* in *ajukdūps* 'eternity'. As already mentioned in a different context by Weiss (1994: 134-135), *\*ajuki-* could formally be based on an old thematic noun *\*ajuka-* 'time; eternity'. PGmc. *\*ajuka-* would have been borrowed as Middle Proto-Fennic *\*ajeka*, as shown by the following well-known parallels:

Baltic *\*angurias* 'eel' → Fi. *ankerias*

Baltic *\*perkūnas* 'oak; the thunder god' → *perkele*, a swear-word

and subsequent deletion of the *\*-e-* as in:

Middle Proto-Fennic *\*kojera* > Fi. *koira* /*kojra*/ 'dog'<sup>1</sup>

We would expect the Germanic noun not only to have meant 'eternity, vitality', but also 'life, age', a double meaning detectable from comparison with other Indo-European languages. Cf. also the meaning '(n)ever' of *\*h<sub>2</sub>óju k<sup>w</sup>id* 'what(ever) time'. contexts contexts.

### 4 Conclusion

The course of events may be summarized as follows:

<sup>1</sup> I thank Petri Kallio for pointing out to me the chronology of these developments.

- a) PIE *\*h<sub>2</sub>ói-u-g(ʷ)-o-* became PGmc. *\*ajuka-*
- b) PGmc. *\*ajuka-* ‘time’ was borrowed to Middle Proto-Fennic and regularly substituted as *\*ajeka*
- c) PGmc. *\*ajuka-* forms the basis of the adjective *\*ajuki-* ‘eternal’ (may also have happened before step a) )
- d) Middle Proto-Fennic *\*ajeka* regularly becomes Late Proto-Fennic *\*ajka*

The Fennic meaning may in any case have been affected by contamination with the inherited lexeme *ikä* ‘age; life; lifetime’, as has been suggested to me by Michael Fortescue (p.c.).

## References

- Aikio 2014 = Luobbal Sámmol Sámmol Ánte (Ante Aikio): “Studies in Uralic Etymology II: Finnic etymologies”. – *Linguistica Uralica* L. 2014, 1: 1-19.
- Andersen, Henning: *Reconstructing Prehistorical Dialects: Initial Vowels in Slavic and Baltic* [= *Trends in Linguistics Studies and Monographs* 91]. Berlin / New York 1996.
- Junttila, Santeri, 2012: “The prehistoric context of the oldest contacts between Baltic and Finnic languages”. – Riho Grünthal & Petri Kallio (eds.): *A Linguistic Map of Prehistoric Northern Europe* [= *Mémoires de la Société Finno-Ougrienne* 266]. Helsinki: Suomalais-Ugrilainen Seura. Pp. 261-296.
- Koivulehto, Jorma, 1973: “Germanisch-Finnische Lehnbeziehungen III”. – *Neuphilologische Mitteilungen* 74: 561-609 [reprinted with postscript in *Verba Mutuata*. Helsinki: Suomalais-Ugrilainen Seura, 1999: 83-120.
- Larsson, Jenny Helena, forthc.: “The lexicon of Baltic”. – Brian Joseph, Joseph Klein & Matthias Fritz (eds.): *Comparative Indo-European Linguistics. An International Handbook of Language Comparison and the Reconstruction of Indo-European*. Berlin: Mouton de Gruyter.
- Liukkonen, Kari, 1999: *Baltisches im Finnischen* [= SUST 235]. Helsinki: Suomalais-Ugrilainen Seura.
- Melchert, H. Craig, 2007: “Luvian Evidence for PIE *\*h<sub>3</sub>eit-* ‘take along; fetch’”. – *Indo-European Studies Bulletin, UCLA* 12, 1: 1-3.
- Tichy, Eva, 2004: “Gr. οἶσθα, lat. *ūtor* und die Mittelteile der Duenos-Inschrift. – *Glotta* 78: 179-202.
- Weiss, Michael, 1994: “Life everlasting: Latin *iūgis* ‘everflowing’, Greek ὑγιής ‘healthy’, Gothic *ajukdūps* ‘eternity’ and Avestan *yauuaēji* ‘living forever’”. – *Münchener Studien zur Sprachwissenschaft* 55: 131-156.



# Albanian *hundë* ‘nose’, and Faroese, SW Norwegian *skon* Finnish *kuono* ‘snout’<sup>1</sup>

## Abstract

Alb. *hundë* ‘nose’ has no accepted etymology, but the stem corresponds regularly to Far. *skon*, Nw. dial. *skon* ‘snout’ if these go back to PGmc. \**skuna-*. Mod.Icel. *skon(n)r*, a fem. nick-name, seems to rule out the alternative reconstruction \**skanō-*. The Balto-Fennic root reflected in Fi. *kuono*, Est. *koon* ‘snout’ is problematic in the light of the vocalism (the only parallel of \**ō* substituting PGmc. \*-*u-* is *ruoste*, *ruosma* ‘rust’), but may nonetheless be a Germanic loan. Former proposals deriving *hundë* from PIE \**skeu-* ‘spring forward’ turn out to be correct, but a derivative \**sku-n-* ‘projection’ (also > Alb. *hu* ‘penis’) must have been formed already before the emergence of Albanian. The suffixal part, PALb. \*-*tā* > -*dë*, either reflects an inner-Alb. formation or goes back to the PIE root extension known from PGmc. \**skundjan-*, \**skundōn-* ‘drive forward’.

## 1 Albanian evidence for PIE reconstruction

Among Indo-Europeanists today, Albanian has acquired, at least unofficially, a discredited reputation as the more or less “useless” Indo-European language branch: it has allegedly retained all too little of the original lexicon, having replaced many everyday words with borrowings from (especially) Slavic, Greek, Latin, and Romance; many other innovations belong to the notoriously shadowy “Ancient Balkan” vocabulary; Albanian often exhibits odd phonological irregularities and aber-

---

1 The present article was published as Hyllested 2012. Apart from this footnote (including the reference just mentioned), the abstract and the exact title, the two articles are identical.

rant derivational patterns; and Albanian lexemes in general are so short that “anything goes” in etymology, provided your semantic analysis is creative enough. In many cases, Albanian forms are mentioned merely to show the geographical representation of a given lexeme.

This reputation, I would argue, is based on a skewed perception of the actual state of affairs. It is true that we know comparatively little about the history of Albanian from its split from the Indo-European core until its earliest attestation, and the internal history of a language branch is indeed more difficult to uncover when the branch in question has few members or only a single one – Proto-Albanian must be reconstructed by dialect material and internal reconstruction (disregarding the poorly attested candidates for close relatives, such as Messapian). But Albanian has potential: a fair description would be to say that much of the lexicon can be defined not as obscure loanwords, but rather as unexplained. In other words, unheeded archaisms might be hiding even in the basic vocabulary, waiting for us to give it another try, applying today’s broader knowledge of PIE and the individual branches. One way forward, which is becoming more and more widespread, seems to be the inclusion, in Indo-European etymology, of evidence from even more peripheral languages and dialects. Peripheral, to be sure; but as it turns out, they may still be extremely relevant and provide crucial information about details in PIE reconstruction. This is at least the case with minor Eastern Iranian languages and Nuristani languages, and in recent years this has proven to be true not least of all in the case of Germanic dialects (cf. Kroonen forthcoming). This article is an attempt to solve an obnoxious etymological riddle by straightforward comparison of forms in peripheral languages.

## 2 Albanian *hundë* ‘nose’; previous proposals

It is indeed a nuisance for comparativists when lexemes from the very core of the basic vocabulary defy a generally accepted etymology. One good example is Albanian *hundë* ‘nose’. As a term for a body part, it belongs to the basic vocabulary par excellence. Hamp (1965: 130) hinted at the initial *h-* as a potential clue to discovery: “It belongs to a list of very basic Albanian words in *h-* which have consistently resisted etymology”.

Since a well-known source of Albanian *h-* is PIE *\*sk-*, scholars have tried to search for possibilities among established roots in PIE or just lexemes in other Indo-European languages with this initial sequence.

Meyer (1891: 153-154) was perhaps on the right track by relating it to PALb. *\*skuna-* > Alb. *hu* ‘penis’ (Gheg *hû*, gen. *hûni*, Tosk *huri* ‘pole; limb’), Gk. *κύνδαλος* ‘peg’, and Skt. *skúndate* ‘spring forth’. Schmidt (1930: 19) accepted Meyer’s etymology right away, and it is also embraced by Ölberg (1972) and reappraised by Orel (1998: 152), who reconstructed for *hundë* a prestige *\*skun-tā*, derived from *\*skuna-*. However, this classical proposal has not won general acceptance because it suffers from the fact that some of the suggested cognates themselves have disputed etymologies. Beekes (2010: 803) regards Gk. *κύνδαλος* as a substratum word.<sup>2</sup>

Pedersen (see Tagliavini 1937: 276) equated *hu* with Lith. *skuja* ‘pine needle’ instead, leaving *hundë* as simply unetymologized. Barić (1919) related *hundë* to Lat. *sentio* ‘to feel’ (via the meanings ‘sense’ and ‘smell’). Çabej (1953) connected it with Rum. dial. *hudă*, *hudră* ‘hole, crack’ (which would then be an “Ancient Balkan” loan of IE origin) and Skt. *kuhara-* m. ‘cavity’, suggesting an original meaning ‘nostril’. Hamp (1965:126) also prefers a development via ‘nostrils’, originally ‘entrances’, from a Proto-Alb. *\*hun-* meaning ‘(an) insert’, the base of the modern verb *hyj*, Gheg *hýj* ‘to enter’ (of disputed origin). He rejects an otherwise formally possible derivative *\*skud-V-nt-* ‘projection’ (< *\*skeud-* ‘shoot’) because the Arvanitic dialect of Sophikó (Greece) has a crucial short vowel that rules out such a contraction. Huld (1984) envisages an Old Alb. metathesis of the PIE ‘nose’ word: *\*noh-* > *\*hon-* + the suffix *-tā*, ultimately from PIE *\*nas-* ‘nose’; but this would involve both suffixation, the as yet shady raising of *\*o* > *u* before nasal (perhaps in *bungë* f., pl. *-a*, ‘chestnut oak, *Quercus sessiliflora*, *Quercus petraea*’)<sup>3</sup> and a metathesis which may seem ad hoc, although not unprecedented: cf. the fact that Alb. *hobe* ‘catapult, sling’ seems to be a development of an originally dialectal *bahe* ‘id.’, a singularized plural of the archaic sg. *bahë*, a loanword from Common Slavic *\*bojъ* ‘fight’ (Orel 1998: 14, 150).

<sup>2</sup> Cf. Kümmel (2010) on the relatively poor occurrence of safe PIE roots containing the sequence -ND-.

<sup>3</sup> According to Orel (1998: 42), Hamp’s reconstruction *bungë* < *\*bug-nā* finds support in *peng* ‘security, pledge’ < Lat. *pignus* (also Demiraj 1997: 112-113), but this is not true if Lat. /gn/ was pronounced [ɲn] (see, e.g., Meiser 1998: 52, 121 on the details). In that case Alb. -ng- is just the rendering of Lat. -ngn-, written <gn>, with loss of the final nasal (by assimilation) in such a cluster.

### 3 The background of Albanian initial *h*-

Let us begin our own analysis by having a closer look at the different sources for initial *h*- in Albanian. Original *h*- is retained in loanwords from Slavic (e.g., Alb. *hitas* ‘to hurry’ < Common Slavic *\*xytati*, Alb. *hukas* ‘to shout’ < Common Slavic *\*hukati*)<sup>4</sup> and Latin (e.g., Alb. *herë* ‘moment of time, hour’ < Lat. *hōra*). But in the case of *hundë*, there is simply no obvious candidate from any neighboring language (disregarding the aforementioned proposal by Çabej).

It cannot be excluded that Alb. initial *h*- sometimes reflects a PIE initial laryngeal. Hamp (1965) reconstructed *\*h<sub>4</sub>* > Alb. *h*- (as in *herdhë* ‘testicle’ ~ Gk. ὄρχις id.) while all other laryngeals disappeared. This was heavily criticized by Ölberg (1972) and has never been widely accepted, although it is accepted by Mallory and Adams (1997: 10); and according to Kortlandt (1998), *\*h<sub>2</sub>e-* and *\*h<sub>3</sub>e-* yield Alb. *ha-* (cf. also Demiraj 1997). Alb. *h*- may indeed show up in the position of an original laryngeal, e.g., Alb. (*h*)*ethe* ‘fever’ < PIE *\*h<sub>2</sub>eǵd-s-* and Alb. *hut* ‘empty, deserted’ < PIE *\*h<sub>2</sub>u-tǵo-* (Goth. *aupēis* ‘barren, desolate’). However, the very occurrence of a laryngeal is not assured in the material in question; and even if one insists on initial consonants in all PIE roots, the picture is blurred by a notorious tendency to insert a spontaneous *h*- before initial vowel in Albanian:

(*h*)*ark* ‘bow’ ← Lat. *arcus* id.

(*h*)*armë* ‘weapon’ ← Lat. *arma* id.

*hikërr* ‘sour milk; buckwheat’ ← an inner-Albanian derivative from *ikëj* ‘to run, go (away)’

*hokë* ‘joke, jest’ ← Lat. *iocus* id. (note that secondary *-j-* and *j-* in loanwords is regularly substituted by Alb. *h*-, e.g., *krahinë* ‘region, district’ < SCr. *krajina*: cf. Rasmussen 1985)

(*h*)*urdhe* f.pl. ‘ivy’ < PIE *\*uǵdʰo-* ‘root, wort’ (OE *word* ‘thorn-bush’)

(*h*)*urdhë* ‘pond, pool’ ~ Common Slavic *\*virō* ‘whirlpool’

The rare PIE onset cluster *\*ks-* regularly yields Alb. *h*- (cf. *i/e huaaj* ‘foreign, strange’ ~ Gk. ξένφος id. and *hirrë* f. ‘whey’ ~ Skt. *kṣirā-* ‘milk’). Again, there is no obvious candidate available. We are left, finally, with PIE *\*sk-* and *\*sḱ-*, which merge into Alb. *h*-, at least before a back vowel (also in inlaut; see, e.g., Huld 1984: 149, Matzinger 2006: 78):

<sup>4</sup> See Svane 1992:256-7.

*halë* f. ‘chaff’ < \**skol-ieh*<sub>2</sub> (Goth. *skalja*, Lith. *skeliù*, *skélti* ‘to hew; to split’)  
*harb* ‘rudeness’ ← < \**skor-b<sup>h</sup>o*- (OHG *scarf*, Latv. *skatbs* ‘sharp, harsh’)  
*hedh* ‘throw’ ~ *hudh* ‘hurl’ < \**skeu-d-* ~ \**sku-d-* (ON *skjóta*) ‘shoot; spring forward’  
*helm* ‘poison’; disputed, but probably connected to OHG *scalmo* ‘plague’, W *claf* ‘sick’ (cf. Hyllested 2010: 111-112)  
*hënë*, Gheg *hânë* ‘moon’ < \**skand-* (Skt. *cándra-* ‘moon’, *ścand-* ‘shine’, Bret. *cann* ‘moon’; Lat. *candeō* ‘glow’, Gk. *κανδαρός* ‘ember’)  
*hi*, def. *hiri*, Tosk dial. *hī*, Gheg *hī*, def. *hīni* ‘ash’ < \*(s)*keniso-* (Lat. *cinis*, gen. *cineris* id., Gk. *κόνις* ‘dust’, Toch. B *kentse* ‘rust’ [sic] < \**koniso-*)<sup>5</sup>  
*hije* f. ‘shadow’ < \**sk<sup>h</sup>h<sub>2</sub>i-eh*<sub>2</sub> (Gk. *σκιά*, Toch. B *skiyo*) ~ \**skéh<sub>2</sub>i-eh*<sub>2</sub> (Skt. *chāyā*)  
*humb* ‘to leave, to lose, to spoil, to miss’ < \**sku-m-b<sup>h</sup>-*, nasal present to ← \**skeu<sup>b</sup>h-* (Lith. *skumbù*, *skùbti* ‘to hurry, to hasten’, Goth. *af-skiuban* ‘to push away, to reject’)  
*hurdhë*, *hudhër* ‘garlic’ ~ Gk. *σκόροδον* id.

#### 4 A new proposal

It seems to have escaped everyone’s attention that several West Nordic languages possess words with almost identical semantics as well as a strong formal similarity to Meyer’s reconstruction: Faroese *skon* f., -ar, -

<sup>5</sup> With PIE \**e* > Alb. *i* either by umlaut from -*i*- in the following syllable (Orel 2000: 145, de Vaan 2004: 70-71) or in secondary consonant clusters such as oblique forms of *s*-stems (which are later contracted); cf., e.g., *vit* ‘year’ < \**uetso-* < \**ueteso-* (Hamp 1971: 121-122). Meyer’s (1891: 152) reconstruction \**skino-*, accepted by Tagliavini (1937: 312), Huld (1994: 74), and Orel (2000: 131, 218), renders impossible the otherwise almost universally accepted equation with Gk. *κόνις* (with *o*-grade); Lat. *cinis* does not reflect original *i*-vocalism, but results from a vowel assimilation \**keni-* > /kini-/, as in *similis* (< \**semilis*) ‘similar, like’. Alternatively, if one prefers to avoid *s-mobile* on the basis of Albanian only, one could derive *hi* from Early Proto-Alb. \**skūja* (~ Gmc. \**skeuja* ‘cloud’, Eng. *sky*) with Early Proto-Alb. *ū* > *i* (~ *y*) preceding -(C)*j*- (Orel 2000: 11-12, cf. *shi* ‘rain’ < \**sūja* ~ OPr. *suge* /su:je/ id., and *mizë*, *myzë* ‘a fly’, with dimin. suff. -*zë*, ~ ON *mý* id. < PGmc. \**mūja-*); but then, in return, one would have to accept secondary nasalization of the vowel as in Gheg *sý* ‘eye’, *drû* ‘wood’.



*ir* ‘nose; snout; face; (colloqu.) mug’ has a counterpart in SW Nw. dial. (Hardanger, Vossemlå, Sogn, Nordhordland; cf. *Grunnmanuskriptet*<sup>6</sup>) *skon* ‘snout, muzzle’, but is otherwise a hitherto obscure and isolated word, not even represented in ON, let alone East Nordic or the rest of Germanic. There are two possible Germanic proto-forms:<sup>7</sup>

- a) ~ ON *\*skōn* < *\*skanō-* (like Far. *lon* f. -ar, -ir ‘longhouse’ < ON *lōn* [~ NE *lane*] or *mon* m. < ON *mōn* ‘mane’, or
- b) ~ ON *\*skon* with *a*-umlaut of PGmc. *\*-u-* < PGmc. *\*skuna-*, *\*skunō*

The latter finds support in the Mod. Icel. fem. nickname *skon(n)r*, because the Norwegian dialect of Vik i Sogn has a similar *skon* ‘hag, poor woman’ (Blöndal 1989), and the alternative would result in Mod. Icel. *†skön(-)*.

## 5 Finnish *kuono* and its relatives

Furthermore, Jorma Koivulehto suggested in 1982 (see Kylstra et al. 1991-2012, II: 125-126) that the Germanic proto-form of *skon* (or an early reflex thereof) had been borrowed into Balto-Fennic, cf. Fi. *kuono*

<sup>6</sup> *Grunnmanuskriptet* (Norwegian ‘Basic manuscript’) is a manuscript dictionary from the 1930’s (completed 1940) which was meant to be released as the first major dictionary of Norwegian in which both lemma and definition were given in Nynorsk. Entries were taken from other monumental dictionaries of Norwegian from that period by Ivar Aasen, Hans Ross, Steinar Schøtt, and others, but it does not contain all dialectal forms from minor dialectal dictionaries, which were meant to be added later. For various reasons, the manuscript was never published as a complete dictionary; but *Dokumentationsprosjektet* (‘The Documentation Project’) at the University of Oslo has made the entire manuscript available in searchable electronic form (see the References below). It now serves as one of the basic sources for the 12-volume *Norsk Ordbok* (‘Norwegian Dictionary’), edited at the University of Oslo, which is projected to be finished in 2014.

<sup>7</sup> PGmc. *\*skēnō-* is not an option since this would yield Early ON *\*skōn*. In Late ON, *ō* merges with *á*, but not if it was nasalized, in which case it (often) merges with *ó*. Cf. for the whole train of events Proto-Norse *\*nahtu* > Early ON *nōtt* > Late ON, Mod. Icel. *nótt*. In Faroese too, nasalized *ō* would give *ó*, cf. ON *vōn*, pl. *vánir* > Far. *vón*, pl. *vánir*. Thus Far. *skon* with short -o- and the pl. *skonir* precludes a reconstruction *\*skēnō-*. The same is true for Nw. *skon*, as *Grunnmanuskriptet* cites the word with the vowel *ō*, which can go back to either ON *ō* (< PGmc. *\*a*) or *o* (< PGmc. *\*u*), but not to ON *á* or *ó*.

‘snout, muzzle’, Est. *koon*, Votic *kōn* id. These forms together point to a proto-form *\*kōnV*, with secondary lengthening of PGmc. short *\*u*, as in Fi. *ruoste* (Est. *rooste*) and Karel. *ruosma* ‘rust’ (< *\*Balto-Fennic \*rōsteh* and *\*rōsma* respectively) – but this is the only other example of such a lengthening; otherwise PGmc. *\*u* is substituted with a Fennic short vowel, mostly *\*u*, but sometimes *\*o* (see Kylstra & al. 1991-2012, I: xviii); examples:

- \*hurskas* ‘pious, devout’ (only in North Balto-Fennic; Fi., Karel. *hurskas*) ← PGmc. *\*hurskaz*
- \*kulta* ‘gold’ (Fi. *kulta*, Est. *kuld*) ← PGmc. *\*gulþa*
- \*kuningas* ‘king’ (Fi., Est. *kuningas*) ← PGmc. *\*kuningaz*
- \*lukko*, *\*lukku* ‘lock’ (Fi. *lukko*, *lukku*, Est. *lukk*) ← PGmc. *\*lukōn-*
- \*multa* ‘humus, topsoil’ (Fi. *multa*, Est. *muld*) ← PGmc. *\*mulðō-*
- \*murkina* ‘breakfast’ (Fi. *murkina*, Est. *murgin(a)*) ← PGmc. *\*murginaz*
- \*ruyis* ‘rye’ (Fi. *ruis*, Est. *rukis*) ← PGmc. *\*rugiz*
- \*turþeh*, *\*turvas* ‘peat, turf’ (Fi. *turve*, Est. *turvas*, dial. *turv*) ← PGmc. *\*turbaz*, *\*turbz*
- \*kotti* ‘bag; scrotum; uterus; trough, etc.’ (Fi. *kotti*, Est. *kott*) ← PGmc. *\*kuððan-*
- \*porðas* ‘degree, step, level; (pl.) stairs: staircase’ (N only; Fi. *porras*, Veps *pordaz*) ← PGmc. *\*burðaz*
- \*sorta-ða* ‘to oppress’ (as a verb only in North Balto-Fennic; Fi. *sortaa* ‘to oppress’; Est. *sõõrd* ‘clearing; margin of a field’) ← PGmc. *\*sturtjan-*

Secondary lengthening, however, does occur in Germanic loanwords with front vowel, cf. Fi. *viikko* < PGmc. *\*wikōn-* ‘week’ and *liikkiö* ‘ham’ ← PGmc. *\*flikkija-*, and there is no doubt that *ruoste* and *ruosma* must be substitutions for PGmc. *\*ruste-* and *\*rusma-*, respectively. The three different outcomes are probably dependent on chronology rather than phonological conditioning. Thus, it cannot be ruled out that *\*kuono* comes from *\*skuna-*.

Nikkilä (see Kylstra & al. 1991-2012, II: 125-126) connects the Fennic word with PGmc. *\*gōnō* (Mod. Icel. *góna* ‘snout, esp. of seal, shark, or wolffish’) instead, and this etymology is perhaps to be preferred, also since the *-o* in Finnish often substitutes a Germanic *ō*-stem. This *-o*,

however, can also represent a Balto-Fennic suffix, and is not present in all of the Balto-Fennic languages in either case. Both etymologies are in any case possible.<sup>8</sup> I find Koivulehto's proposal that PGmc. *\*skuna-* (via 'muzzle') also stands behind Karel. *kynä* 'trough', Est. *küna*, Liv. *kinà* (Kylstra et al. 1991-2012, II: 137) rather far-fetched.<sup>9</sup>

## 6 PIE reconstruction: Formal possibilities

Now, a PGmc. *\*skuna-* 'snout' obviously fits well with Proto-Albanian *\*skuntā* 'nose', even if there are different ways of accounting for the Albanian *\*-t-*. Considering the range of meanings of *hundë*, it would make sense to maintain relationship with the 'shoot' root:

- a) 'nose'
- b) 'point, tip, summit'
- c) 'projection, overhang'
- d) 'promontory, headland, cape'

Affinity with Lith. *skutnà* 'bald head' is formally possible since its expected Albanian counterpart would be exactly *\*hundë* (< *\*skuntā* < *\*skutnā*; cf. the metathesis in *bung(ë)* 'chestnut oak' described above in 2.). Since the Lith. forms and the underlying verb *skūsti* 'to shave, to peel' are most likely connected to MIr. *scoth* f. 'point, edge' < PCelt. *\*skutā*, it does not appear semantically impossible either. Note also the Hesychian gloss σκῦτη, of similar shape, which is given the meaning κεφαλή 'head'.

It is also conceivable, however, that *\*skuntā* simply reflects the original order of nasal and stop. We know Germanic forms like OHG *scunten*, OE *scyndan*, ON *skunda* → Eng. *scoon* (→ *schooner*), *scun* 'fly forward' < PGmc. *\*skundjan-*, *\*skundōn-* 'drive forward', and these either go back to PIE *\*skundh-* or Verner variants of *\*skunt-*; when compared to No. dial. *skut* m. 'projection, overhang', ON *skúta* 'kind of ship';

<sup>8</sup> Paul Kiparsky (p. c.) has reminded me that original mid vowels are lengthened in open syllables in Fennic languages; but as a regular development this takes place much earlier, on the way from Uralic to Fenno-Ugric or Fenno-Permian, and would hardly affect *\*-u-* in Germanic loanwords in Balto-Fennic, let alone a later Gmc. *\*-o-* after the operation of *a*-umlaut.

<sup>9</sup> As an alternative, Skt. *kuṇḍa-* 'hole in the ground, pit, etc.' comes to mind, but this is a borrowing from Dravidian.

schooner' (→ OIr. *scúta*, NE *scout*, MDu. *scūte*), Lith. *skudrūs* 'fast', Skt. *códati* 'drive forward', it is clear that we must assert at least two root-variants with different stops, hence probably old extensions of an original root *\*skeu-*. Nasalized forms may represent generalized nasal presents. Alb. *hundë* cannot reflect a variant with a voiced stop, which would be lost in the position after a nasal, yielding †*hunë*.

I conclude that *hundë* is ultimately related to Alb. *hu* 'penis', *hedh* 'to throw', *hudh* 'to hurl', *hyj* 'to enter' and *humb* 'to leave; to lose; to spoil; to miss' as originally suggested by Meyer (and followed by Schmidt, Ölberg, Orel, and partly Hamp), but this word-family cannot be safely established by internal reconstruction alone. It is Germanic *\*skuna-*, reconstructed on the basis of Faroese and Norwegian material, as well as possible ancient Germanic forms in Balto-Fennic, that have provided the clue.

Since a primary word for 'nose' is already known from most Indo-European languages, and since this word is known to be at least of PIE age (PIE *\*nas-*),<sup>10</sup> it seems reasonable to reconstruct the meaning of *\*skun-o-*, *\*skun-to-* rather as 'snout' (i.e., 'nose of an animal [as opposed to the human nose]'), preserved in Germanic and having replaced the original 'nose' word in Albanian only. Thus *\*skun-o-*, *\*skun-to-* would be of at least Northwest Indo-European age.

## References

- Barić, Henrik, 1919: *Albano-rumänische Studien I* [= *Zur Kunde der Balkanhalbinsel. Quellen und Forschungen VII*]. Sarajevo: Institut für Balkanforschung.
- Beekes, Robert, 2010: *Etymological Dictionary of Greek*. Leiden & Boston: Brill.
- Çabej, Eqrem, 1953: "Grupet *nd, ng në* gjuhën shqipe". – *Buletin për Shkencat Shoqërore* 4: 30-38. Tirana.
- Demiraj, Bardhyl, 1997: *Albanische Etymologien*. Amsterdam: Rodopi.
- Grunnmanuskriptet*.  
[http://www.dokpro.uio.no/nynorsk/gmanus/gmanussoek\\_fside.html](http://www.dokpro.uio.no/nynorsk/gmanus/gmanussoek_fside.html) (accessed 11 July, 2012).
- Hamp, Eric P., 1965: "Evidence in Albanian". – Werner Winter (ed.), *Evidence for Laryngeals*.
- Hamp, Eric P., 1971: "'Fils' et 'fille' en italique: nouvelle contribution". – *Bulletin de la Société Linguistique de Paris* 66: 213-27.
- Huld, Martin E., 1984: *Basic Albanian Etymologies*. Columbus, OH: Slavica.

<sup>10</sup> Details of the reconstruction vary, but everyone agrees about the existence of the etymon.

- Huld, Martin E., 1998: "Meillet's Northwest Indo-European Revisited". – Karlene Jones-Bley & Martin E. Huld: (eds.), *The Indo-Europeanization of Northern Europe* [= *Journal of Indo-European Studies* Monograph 17]. Washington: Institute for the Study of Man. Pp. 109-125.
- Hyllested, Adam, 2010: "The Precursors of Celtic and Germanic". – Stephanie Jamison, H. Craig Melchert, and Brent Vine (eds.): *Proceedings of the 21<sup>st</sup> Annual UCLA Indo-European Conference*. Bremen: Hempen. Pp. 107-128.
- Hyllested, Adam, 2012: "Albanian *skon* 'nose' and Faroese, SW Norwegian *skon* 'snout'". – Stephanie Jamison, H. Craig Melchert, and Brent Vine (eds.): *Proceedings of the 23<sup>rd</sup> Annual UCLA Indo-European Conference*. Bremen: Hempen. Pp. 73-81.
- Koivulehto, Jorma, 1982: "Zur Periodisierung germanischer Lehnwörter im Ostseefinnischen: Wie findet man germanische Lehnwörter und was sagen sie uns?". Lecture given at Rijksuniversiteit Groningen, 12 May 1982.
- Kortlandt, Frederik, 1998: "Reflexes of Indo-European Consonants in Albanian". – *Orpheus* 8: 35-37 [Reprinted in F. K., *Studies in Germanic, Indo-European and Indo-Uralic*, 329-332]. Amsterdam: Rodopi, 2010.
- Kroonen, Guus, forthcoming: *Etymological Dictionary of Proto-Germanic*. Leiden: Brill.
- Kümmel, Martin, 2010: "The Distribution of Roots Ending in IE \*ND". Paper presented at *The Sound of Indo-European 2* (Opava, November 2010). Forthcoming in the proceedings volume, Roman Sukač et al. (eds.).
- Kylstra, A. D., Sirkka-Liisa Hahmo, Tette Hofstra, and Osmo Nikkilä, 1991-2012. *Lexikon der älteren germanischen Lehnwörter in den ostseefinnischen Sprachen I-III*. Amsterdam / Atlanta: Rodopi.
- Magnússon, Asgeir Blöndal, 1989: *Íslensk Orðsifjabók*. Reykjavík: Orðabók Háskólans.
- Mallory, James P., and Douglas Q. Adams, 1997: *Encyclopedia of Indo-European Culture*. London / Chicago: Fitzroy Dearborn.
- Matzinger, Joachim, 2006: *Der Altalbanische Text Mbsuame e Krështerë (Dottrina cristiana) des Lekë Matrënga von 1592: Eine Einführung in die albanische Sprachwissenschaft*. Dettelbach: Röhl.
- Meiser, Gerhard, 1998: *Historische Laut- und Formenlehre der lateinischen Sprache*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Meyer, Gustav, 1891: *Etymologisches Wörterbuch der Albanesischen Sprache*. Strassburg: Trübner.
- Oettinger, Norbert, 2003: "Neuerungen in Lexikon und Wortbildung des Nordwestindogermanischen". – Alfred Bammesberger & Theo Vennemann (eds.): *Languages in Prehistoric Europe*. Heidelberg: Winter. Pp. 183-193.
- Ölberg, Hermann, 1972: "Einige Überlegungen zur Laryngalthorie. An Hand des Albanischen". – *KZ* 86: 121-136.
- Orel, Vladimir, 1985: Review of Huld 1984. – *Ėtimologija* 179-182.
- Orel, Vladimir, 1998: *Albanian Etymological Dictionary*. Leiden / Boston: Brill.
- Orel, Vladimir, 2000. *A Concise Historical Grammar of the Albanian Language*. Leiden / Boston: Brill.

- Rasmussen, Jens Elmegård, 1985: "Miscellaneous morphological problems in Indo-European languages I-II". – *Lingua Posnaniensis* 28: 27-62 [Reprinted in J.E. Rasmussen, 1999: *Selected Papers in Indo-European Linguistics*. Copenhagen: Museum Tusculanum. Pp. 100-144.
- Schmidt, Manfred Erwin, 1930: "Untersuchungen zur albanischen Lautgeschichte". – *KZ* 57: 1-42.
- Svane, Gunnar, 1992: *Die slavische Lehnwörter im Albanischen*. Aarhus: Aarhus University Press.
- Tagliavini, Carlo, 1937: *L'albanese di Dalmazia: Contributi alla conoscenza del dialetto ghegho di Borgo Erizzo presso Zara*. Firenze: Olschki.
- de Vaan, Michiel, 2004: "PIE \*e in Albanian". – *Die Sprache* 44, 1: 70-85. Stockholm / Gothenburg / Uppsala: Almquist & Wiksell.
- Wiklund, K.B., 1917-1920: "Die ältesten germanischen Lehnwörter im finnischen. – *Indogermanische Forschungen* 38: 48-114. (61-62)
- Witczak, Krzysztof Tomasz, 2003: *Indoeuropejskie nazwy zbóż*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
- Witczak, Krzysztof Tomasz, 2004: "Albanian *kophst* 'garden' and its cognates". – *Studia Etymologica Cracoviensia* 9:179-787.



## ***Estonia and the Aestii:*** **Baltic Etymology as a Key to Fennic Ethnonyms**

### **Abstract**

The name of the Estonians, Est. *Eesti*, goes back to the tribal name \**Aestii* first mentioned by Tacitus, denoting inhabitants of the Baltic in the broadest sense. Several sources reveal that the name originally had an -r- in the stem. That the original diphthong was \**ai-* as in the Latinized forms is assured by the Old Gutnish attestation *Aistland*. The name must come from Baltic \**aistra-* ‘pimple (m.); vehement passion (f.)’ which can be reconstructed on the basis of a) a loanword in Livonian *aistar* ‘pimple’ and its variant *vistar*; b) Lith. *aistrà* and c) cognates in other Indo-European languages with the same double meaning. The name was given to the Estonians as a translation of PGmc. \**finōn-*. Numerous parallelsexist among Fenno-Ugric ethnonyms, perhaps pointing to an old designation referring to fish scale as money or fish skin as valuable garment.

### **2 Attestations of the name**

The name of Estonia, Est. *Eesti*, goes back to the tribal name \**Aestii* first mentioned by Tacitus in 98 BC (gen.pl. *Aestiorum gentes* Germ. 45, 2) and later by Cassiodorus 523-526 BC (*H(a)estis Theodoricus rex*, *Variae* 5, 2); Jordanes, d. 552 (gen.pl. *Aestorum natio*, *Get.* 23, 119); Einhard in 830 (*Aisti* in *Vita Karoli Magni*); Wulfstan in 890 (*to, mid Ēstum* in his travel account); and Adam of Bremen in 1073 (*Haisti* and *Aestland*, *Hamburgische Kirchengeschichte*, 12 and §17). On the basis of these attestations, we can identify a Latinized ethnonym *Aestii* or *Aesti* and a stem *Aest-* (*Aist-*).

According to A. Bammesberger & S. Karaliūnas (1998), rather than denoting a specifically Baltic tribe, it was probably a cover term for all



inhabitants of the East coast, including those of Balto-Fennic origin. They point to the fact that, even if the *Aestii* of the antique and early medieval sources denoted Southern Baltic tribes (speaking Baltic languages), the ethnonym has been used by Scandinavians since the Viking Age to denote the Fenno-Ugric forefathers of the contemporary Estonians. Since the 10<sup>th</sup> century the country and their inhabitants were called *Eistr*, more rarely *Eistir* and *Eistland*.

That the original diphthong was *\*ai-* as in the Latinized forms (the only option if the name is inherited into Old Norse, but not if it is late loan) is assured by the Old Gutnish attestation *utan foru i aina oy viþr Aistland, sum haitir Dagaiþi* '[they] travelled to an island off Estonia called Dagö' (Guta Saga, 13<sup>th</sup> c.).<sup>1</sup> Furthermore, *-ai-* occurs in ethnonymic compound *Tafaistr* 'inhabitant of the Häme region in Finland', used as a personal name on an 11<sup>th</sup> c. runestone (U 722) from Löts parish, Trödg's Härad in the Swedish region of Uppland: **tafaistr · lit · raisa · stain · at · a[----- b]roþur · sin** 'Tafeistr raised this stone in memory of his brother'. Otherwise this name occurs with *-ei-*, although the modern Swedish term is *tavaster* (cf. also *Tavastehus* 'the Finnish province of Häme').

## 2 Etymology

Bammesberger & Karaliūnas (1998: 47) further establish on a balance of probabilities that the name originally meant '(agricultural) land; acre' and is derived from Proto-Baltic *\*āist-* 'to burn'<sup>2</sup> either via a meaning 'dry land' (as opposed to the sea), or, more plausibly, referring to the use of slash-and-burn cultivation techniques. They detect this stem also in Baltic place-names like *Aistere* and *Aisternīki* in Latvia and *Eistrai* in Lithuania (with secondary *Ei-* for *\*Ai-*), the typically Latvian suffix *-ere* of the former showing that the name is indeed of Baltic origin; and they compare e.g. Lit. *dogmė* 'glade; acreage, strip of land' ~ *dėgti* 'to burn'; Lat. *terrā* 'land' < *\*ters-* 'dry'. Their main single argument is that other names of Baltic and Balto-Fennic tribes are etymologized as 'land', and that the traditional Estonian self-designation *maarahvas* still in the modern language is literally 'land-people', *maakeel* 'Estonian language'

<sup>1</sup> I thank Sean Vrieland (p.c.) for having pointed this out to me.

<sup>2</sup> Originated from *\*h<sub>2</sub>eidh-t-* as in Lat. *aestus* 'heat', *aestas* 'summer' and PGmc. *\*h<sub>2</sub>eidh-* 'to burn'. For an overview of earlier etymological proposals, see Bammesberger & Karaliūnas 1998: 46, fn. 1.

correspondingly lit. ‘land-tongue’, *maa-sõna* ‘genuine Estonian expression’, lit. ‘land-word’, etc. Since this term is synchronically analyzable it is of course a late variant which may nonetheless preserve semantics and connotations from earlier ethnonyms.

What Bammesberger & Karaliūnas fail to mention is that this term contrasted specifically with the name of another Balto-Fennic population, the closely related Livonians, who were primarily fishermen and literally called themselves ‘inhabitants of the coast’, *rāndalist*, who spoke *rāndakēļ* ‘Livonian’, lit. ‘coast(al) tongue’.

As the authors note themselves, there are indications that the original stem contained an *-r-*: The Old Norse *Ynglingatal* talks about *Eistra dolgi* ‘the Estonian enemy’, and in *Historia Norvegiae*, the name shows up as *Eistriam*, *Eistriis* (in the passage *devenit in Eistriam, puer Olavis Eistriis in servum venundatur* ‘came to Eistria and bought the boy Olav from Eistria as a slave’). They seem to interpret this as an indication that the ethnonym was borrowed from the suffixated form occurring in the place-name *Aīstere*. However, the suffix *-ere* is almost certainly not common Baltic; it is typical of Latvian place-names only and specifically names connected to lakes. Būga (1923: 383) simply considers it a borrowing from Balto-Fennic *\*järvi* ‘lake’ (Fi. *järvi*; cf. also Balode & Bušs 2007: 37).

If the basis for the full sequence *\*aistr-* is common Baltic, there are other possibilities, however. Asserting *-r-* as an original part of the stem provides a formal basis for comparing it to Lith. *aistrà* ‘intense passion’ and Livonian *àistar* ‘pimple, blister; maggot or worm in animal skin; cockchafer grub’. Since Thomsen (1890) the latter has been considered a loan from the Baltic word behind Lith. *ankštirai*, *inkstīras* ‘pimple’, Latv. *ankstīras* ‘larva under the skin of cattle’. However, there are several structural problems implied in this etymology. The expected form is rather something like *†ahtar* since Baltic *\*(n)kš-* normally is substituted with *-h-* in Balto-Fennic, cf. e.g. *ahingas* ‘fish-spear, rake’ ~ Lith. *ākštinās*. One further issue is the Livonian dialectal variant *vistar* which is even further away from the alleged Baltic point of departure.

### 3 Alternative Etymology: Baltic *\*aistra-* is a translation of PGmc. *\*finōn*

I suggest instead that the Livonian word is borrowed from a hypothetical Proto-Baltic *\*aistra-* whose feminine-collective counterpart is at-

tested as Lith. *aistrà*. The former would make up a form identical even in gender to Gk. οἷστρος ‘gadfly; intense passion’, otherwise missing among the IE cognates, and both meanings of the Greek word would then also be covered by Baltic. In Late (Western) PIE *\*oṡstro-* probably arose as a merger of two originally independent words:

- a) *\*(h<sub>2</sub>)oid-tro-* ‘one who makes (sth) swell’ from *\*h<sub>2</sub>eṡd-* or *\*aṡd-* ‘swell’ (cf. *-d-* in e.g. Latv. *idra* ‘disease mark on tree’, Gk. *oṡdos* <οἰδος> ‘tumour’, οἰδέω ‘to swell’, PGmc. *\*aita-* ‘abscess, ulcer’, *\*aitra-* ‘pus’, Lat. *aemidus*).
- b) *\*h<sub>3</sub>eṡs-tro-* (*\*h<sub>3</sub>oṡs-tro-*) ‘irritator, one who sets (sby) in vehement motion’ from *\*h<sub>3</sub>eṡs-* ‘set in vehement motion, urge, irritate’<sup>3</sup> (cf. *\*-s-* in Av. *aešma-* m. ‘anger’, Lat. *īra* ‘id.’ next to Gk. οἶμα ‘rush, attack, rage (of animals)’).

Such a merger could have been facilitated not only by the sudden homonymy (after loss of laryngeal and development of *\*TT > sT*), but even by the semantic connection between stinging insects and swellings on the skin of animals caused by bugs that operate there.

The etymology proposed here also makes it easier to explain the Liv. variant *vīstar* as an East Baltic variant *\*viestra-* with prothetic *\*vie-* as reflex of *\*(H)oi-* as also in *\*vienas* ‘1’ < PIE *\*óih<sub>2</sub>-no-s*.

The comparison can be justified semantically, seeing that many other Fennic ethnonyms bear strong similarities to appellatives with meanings covered by the PGmc. word *\*finōn-*. These meanings are:

- 1) ‘fish fin’ (Sw. *fena*, older *fina*, MLG *vinne*, OE *finn* > Eng. *fin*)
- 2) ‘pimple; abscess’ (Sw. *finne*, NHG *Finne*, Da. *filipens* < finne-pind)
- 3) ‘kornaks’ (Sw. *Dial. fen(a) etc.*)
- 4) Nw. dial. *finn(e)* ‘small horn on animal’, ‘stiff grass species’
- 5) ‘any protuberance on the skin of humans or animals related to diseases, including larvae under the skin, notably on fish (MLG *vinne*, Dutch *vin*, NHG *Finne*)
- 6) ‘finne’ > ‘patch’, cf. NHG *Flosse* ‘fin’ ~ Da. *flosset*

I reconstruct the common Proto-Germanic meaning as ‘protuberance on the skin (including fins and fish scale, larvae under the skin). That

<sup>3</sup> Beekes (2010: 1062) regards *\*h<sub>3</sub>eṡs-tro-* as the only source of οἷστρος.

the meaning 5) of *\*finōn-* matches Baltic *\*aistra-* so well is in itself hardly a coincidence. When compared to homonyms of neighbouring Balto-Fennic tribes, the equivalence becomes even more striking, cf. that Fi. *Suomi* 'Finland' is very reminiscent of *suomu* 'fish scale'. The two are normally considered unrelated, but has been suggested as one possible etymology of the name (cf. Grünthal 1997; Kulonen 1998). Some of the other names have also been viewed as loan translations (Napol'skich 2007): *Lapp* ~ Danish *lap* 'patch', *Veps* ~ Saami *\*veapsē* 'fin', Fi. *Vatja* 'Vote' ~ Balt. *\*vadjā-* 'patch', the ethnonym *sambi*, *sembi* ~ Fi. *sampi* 'sturgeon', which is known to derive from an earlier meaning 'horn core on sturgeon' < 'back-fin of a fish' (Napol'skich 2003).

In South Moravian dialects of Czech, the word *maďar* means both 'Hungarian' and 'pimple', but in this case the primary meaning of *maďar* is clearly 'Hungarian', coming from Hung. *magyar*, meaning that the signification 'pimple' must originate from a wordplay on the double meaning of *uher*. This example thus serves merely as a typological parallel where the semantic shift went in the reverse direction, from ethnonym to appellative.

This situation probably reflects several erroneous translations and folk-etymologies through time, cf. further Old Norse *seiðmenn* 'Saami people', originally 'sorcerers', but synchronically also understandable as 'pollack men'; and ON *taf-eistr* 'inhabitant of Häme', lit. 'tap Estonian' ~ latinized in as *Historia Norvegiae cornuti fenni*, Old Norse *horn-finnar*. Correspondingly, *\*aistra-* might simply be the Baltic translation of what had at some point ended up as the default designation for Balto-Fennic tribes.

#### 4 Other possibilities than mere translations

Although wordplay, erroneous translations and folk-etymology does seem to have played a role for the emergence of some of the ethnonyms, it is on the other hand quite hard to imagine a scenario where these phenomena took place so consistently and repetitively through the ages in different geographical areas, and that they would have resulted in a surviving established ethnonym in every single case. It is thus hardly credible to present the whole story entirely as one of unconscious and incidental calques and misunderstandings. It seems at least worth investigating whether an inherited conceptual association between Fenno-

Ugric peoples and at least one of the meanings involved could be held accountable for some of the homonymies.

Interesting in this context is Widmer's (2003) reconstruction of a Uralic word for 'fish scale', \**kāmV*, which after the beginning of minting (at least after the first half of the 7<sup>th</sup> c. BC) became used as a term for 'coin', especially 'silver coin', 'copper coin', but also 'silver in general', 'silver jewel' and in Ob-Ugric folklore even established as an image of wealth, partly as an attribute to the Urmutter (*kam naj* 'wealth(y) noble-lady'), partly of any material making up status symbols of heroes (weapons, booty). In a version of an Mansi legend, the last word in the formula *śoap̄ar-nē koam-nē* 'silver woman, wealth woman' has been replaced with a word meaning 'Russian', *śāp̄ar-nē kāper-nē*, and in another version, even the first word has been replaced with 'Siberian', *śāp̄ar-nē kööp̄ar-nē*. If some of the Fenno-Ugric peoples had once been named or named themselves after precious metal, it would not be much different from deriving *türk* from the word for 'silk'.

It is also conceivable that a term 'fish-skin' could have constituted an exact parallel to *türk*, referring to an important garment. As noted elsewhere in this dissertation, Fenno-Ugric forest peoples have traditionally produced clothing – coats, boots, and caps – from fish-skin, justifying an analogy with terrestrial animals hunted for their pelt (Armstrong 1997). The burbot along with the sturgeon and sterlet were so important for the Khanty (Ostyaks) in the time of Russian expansions in the 1600s that a particularly bad fishing season could threaten the very existence of a tribe.

## References

- Anttila, Raimo & Sheila Embleton, 1994: "On the Origin of *Suomi* 'Finland'". – *Scandinavian-Canadian Studies* VII : 25-32.
- Balode, Laimute & Ojārs Bušs, 2007: "Latvian place-names of Finno-Ugric origin". – Ritva Liisa Pitkänen & Janne Saarikivi (eds.): *Borrowing of place names in the Uralian languages*. Debrecen / Helsinki: Onomastica Uralica. Pp. 27-44.
- Bammesberger, Alfred & Simas Karaliūnas, 1998: "Zur Frage nach der ethnischen Identität der Aisten". – Alfred Bammesberger (ed.): *Baltistik: Aufgaben und Methoden*. Heidelberg: Carl Winter. Pp. 39-51.
- Būga, Kazimieras, 1923: Review of Jānis Endzelīns & al. (eds.): *Latvijas vietu vārdi*, part 1 (names from Vidzeme). – *Tauta ir žodis* 1: 376-397, 441-444.

- Carpelan, Christian, 1998: "Suomi, häme, sabme sekä finne arkeologian näkökulmasta". – Johanna Laakso & Riho Grünthal (eds.): *Oekeeta asijoo. Commentationes Fenno-Ugricae in honorem Seppo Suhonen sexagenarii* 16.V.1998 [= *Mémoires de la Société Fenno-Ougrienne* 228]. Helsinki: Suomalais-Ugrilainen Seura. Pp. 76-88.
- Décsy, Gyula, 1965: *Einführung in die finnisch-ugrischen Sprachwissenschaft*. Wiesbaden: Otto Harassowitz.
- Décsy, Gyula, 1987: "Lappisch *eadne* und das Suomi-problem". – *Ural-Altaische Jahrbücher* 59 : 45-48.
- Grünthal, Riho, 1997: *Liivistä livviin : Itämerensuomalaiset etnonymit* [= *Castrenianumin toimeitteita* 51]. Helsinki.
- Jämsä, T., 1989: "Finnish *Lappi* 'Lapland' and *lappalainen* 'a lapp'". – *Ural-Altaische Jahrbücher* 61 : 129-132.
- Kallio, Petri, 1998: "Suomi(ttavia etymologioita)". – *Virittäjä* 4: 613-620.
- Katz, M. & Koch, Ch., 1987: "(J)Ugra. Zur ursprünglichen Gestalt des Ugriernamens". *Ural-Altaische Jahrbücher*, N.F. 7: 99-104.
- Koivulehto, Jorma, 1993: "Suomi". – *Virittäjä* 97, 3, 400-407.
- Kulonen, Ulla-Maija, 1998: "Benämningar på finsk-ugriska folk". – *Finsk-ugriska dager 1995 : Foredragssamling om finsk-ugriska emner*. Oslo : Institut for østeuropeiske og orientalske studier / Unipub, Akademika AS. Pp. 23-32.
- Lindroth, H., 1917: "Onomatologiska anmärkningar till T.E. Karstens. Germanisch-finnische Lehnwortstudien". – *Namn och Bygd* 5: 19-52.
- Napol'skikh, Vladimir, 2003: "Uralic fish-names and original home". – *Ural-Altaische Jahrbücher* N.F. 12: 35-57
- Napol'skikh, Vladimir, 2004: "Der Name der Vepsen und frühere Berührungen zwischen Germanen, Lappen und Ostseefinnen". – Peter Anreiter, Marialuise Haslinger & Heinz Dieter Pohl (eds.) *Artes et Scientiae: Festschrift für Ralf-Peter Ritter zum 65. Geburtstag*. Wien: Edition Praesens. Pp. 307-313.
- Parpola, Asko, 2004: "Old Norse *seiðr*, Finnish *seita* and Saami shamanism". – I. Hyvärinen, P. Kallio, and J. Korhonen (eds.), *Etymologie, Entlehnungen und Entwicklungen: Festschrift für Jorma Koivulehto zum 70. Geburtstag* [= *Mémoires de la Société Néophilologique de Helsinki* 63]. Helsinki. 235-273.
- Ross, Alan S.C., 1981 : *The Terfinnas and Beormas of Ohtere*. – London : Viking Society for Northern Research / University College London.
- von Sadovzsky, Otto J., 1995: *Fish, Symbol and Myth: A historical semantic reconstruction* [= *ISTOR books* 6]. Budapest / Los Angeles: Akadémiai Kiadó / International Society for Trans-Oceanic Research.
- Stachowski, Marek, 2002 : "Zur Frage der ursprünglichen Bedeutung von finnisch *Lappi* 'Lapland'". – *Studia Etymologica Cracoviensia* 7: 151-158.
- Thomsen, Vilhelm, 1890 : *Beröringer*
- UEW = Károly Rédei 1988: *Uralisches Etymologisches Wörterbuch* I-III. Budapest : Akadémiai Kiadó.
- Widmer, Anna, 2001: "Ural. \**kāmV* 'Fischschuppe' im Obugr., Sam. und Perm.: zur Benennungsmotivation von 'Münze', 'Metall' und 'Schmuck'. – Heiner Eichner, Peter-Arnold Mumm, Oswald Panagl & Eberhard Winkler (eds.): *Fremd und Eigen : In memoriam Hartmut Katz*. Wien : Edition Praesens. Pp. 279-291.



## Word Migration on the Silk Road: The Etymology of English *silk* and its Congeners<sup>1</sup>

### Abstract

European terms for ‘silk’ display exceptional variation for deriving ultimately from a single Old Chinese source. It is suggested here that ON Norse *silki* ‘silk’, transmitted via the Varangians’ trade in Byzantium through Kievan Rus’ to Scandinavia in the 9<sup>th</sup> c., and OE *seoluc*, OHG *silihho* possibly in a separate wave a couple of centuries earlier, are loanwords from the Iranian language Alanic (Sarmatian) at the Western end of the Silk Road. It reflects a regular Alanic development of *\*ri* > *l*, the original form being *\*sirika-* which leaves several possibilities open: Either a) it is a productive inner-Iranian formation with the suffix *\*(V)ka-*, possibly meaning ‘silk man, inhabitant of Silis’; b) it is a nativization of Gk. *σηρικόν*; or c) it has entered Alanic via Turkic *\*sir(e)-lek* ‘silk garment’ or Mongolian *sirkeg* ‘silk fabric’. Modern Ossetic *zæly*, *zældag*, despite its typical loanword characteristics, may be inherited directly from Alanic *\*silika-*, only reshaped in analogy with *zældæ* (today ‘young grass; turf’) which must have preserved its original meaning ‘golden; yellow’ into medieval Iassic and preserved with that meaning as a loanword in Hungarian *zöld*.

---

<sup>1</sup> This article will be published in Berit Hildebrandt (ed.): *Exchange along the Silk Roads between Rome and China in antiquity: The Silk Trade*. Oxford: Oxbow 2014. Apart from the notations (abbreviations of language names and transliterations of Greek words) the two versions are almost completely identical.



## 1 'Silk' in Europe and Asia<sup>2</sup>

The exchange of silk along the Silk Road is often examined through archaeological finds and the interpretation of historical texts. This contribution seeks to address the question of exchange along the Silk Road and especially the ways through which the silk reached the West by discussing the etymology of *silk* in Western and Northern European languages from antiquity to modern times. The chronological scope of this article is vast in order to better understand the different national concepts through which the material was classified, the different routes the silk trade could take and the ethnic groups that were involved.

Most etymological dictionaries (notably ODEE 827, de Vries 1962: 487, Vasmer 1953-58, III: 387, Falk/Torp 1960: 966-967) agree with and basically just repeat the standard etymology of Eng. *silk*, OE *seoloc*, *seoluc*, *sioloc*, *seolc*, and its immediate congeners in Germanic (ON *silki*, Da., Nw., Sw. *silke*<sup>3</sup>, borrowed into Finnish and Karelian as *silkki* and into the Western Saami languages, cf. SKES IV: 1025<sup>4</sup>; OHG *silehho* 'toga', *selachin* 'cover'), Baltic (Lith. *šilkas*, *šilkaĩ*, Samogitian dial. *silkaĩ*, borrowed into Latvian in the expression *silkuôts* 'sewn with silk'; OPr. *silkas*) and East Slavic (ORu. *šilkŭ*, borrowed into the New Curonian dialect of Latvian as *šilks* and continued in Standard Ru. *šēlk*, Belaru. *šolk*, Ukr. *šovk* and N Ru. *šulk*, from where it has been transmitted to North-Eastern Balto-Fennic languages, and, via Karelian, further into Eastern Saami<sup>5</sup>). The Baltic forms are regarded as old loans from Slavic (Fraenkel 1962-65: 983-984). Schrader (1904-1905: 84) states that the Germanic forms must have been transmitted via Slavic.

<sup>2</sup> I owe my sincere gratitude to Peter Kerkhof, Seán Vrieland, Berit Hildebrandt and two anonymous referees for their insightful comments and invaluable amendments.

<sup>3</sup> Swedish is one of the very few languages that distinguishes two basic words for the silk thread (*silke*) and woven silk (*siden*). The name of the Silk Road is called *Sidenvägen*, thus referring to the fabric.

<sup>4</sup> S Saami *silke*, Pite Saami *silhke*, Lule Saami *silhkē*, N Saami *silki*, Inari Saami *silkke* (SKES IV: 1025).

<sup>5</sup> The Balto-Fennic forms borrowed from Northern Russian are Eastern (Karelian and Ingrian) Fi. *sulkku*, Karelian proper and Olonets *šulkku*, Lude *šulk(u)*, Veps *šūk* (cf. Plöger 1973: 190). From Karelian proper comes Skolt and Kildin Saami *šolkk*, while in Votic, the only Southern Balto-Fennic language to possess the word, *šolkka*, *šonkk* is borrowed directly from Standard Russian. Cf. SKES IV: 1103.

All refer to the ultimate source as Old Chinese, the language of the area where the silk industry began and became important as early as the 3<sup>rd</sup> millennium BC (cf. Wang 1993: 225). Most scholars further envisage that the journey of this word towards the West in some way involves Written Mong.<sup>6</sup> *sirkeg* ‘silk fabric’ and Manchu *sirge*, *sirhe* ‘silk thread, silk floss from a cocoon; string of a musical instrument’, thus pointing to the important role nomadic tribes played in the distribution of silk. Hardly surprising, the standard handbooks also agree that the word was transmitted in antiquity with the trade of silk fabric along the silk road. It is a typical wanderwort of the later, historical kind, where the ultimate origin is at least superficially rather obvious but the ways of transmission less certain (while, for prehistoric culture-words, even the source is often obscure). In this particular case, to our benefit, the transcontinental Silk Road constitutes a concrete historical track on which we can hope to trace the word and its intermediate stations on its way from East Asia to Europe and catch glimpses of the transmission of not only the word, but also the material along the Silk Road.

The Chinese source is the precursor of the Mod.Ch. 絲 *sī* ‘silk; thread; string’; it is commonly reconstructed as OCh. \**sə* or \**siag* (thus Wang 1993 with references) and Middle Ch. \**si*. It is related to other Sino-Tibetan words denoting ‘thread’, ‘string’ or ‘sinew’. Neighbouring Asiatic languages all reflect a final *r*-element, which, if not reconstructable for Old or Middle Chinese, must be explained as suffixal in one of the lending languages from which it can have been transferred further: Middle Kor. *sir* (> Mod. Kor. *shil*), Manchu *sirge*, *sirhe*, Written Mong. *sirkeg*. It is however possible that the *-r-* does go back to Chinese and reflect a second noun 人 *rén* ‘people’ (Genaust 1996: 578) in which case the word borrowed from neighbouring languages would not be a designation for ‘silk’ as such, but rather a compound-like ethnonym already at the time of contact (whose meaning would correspond exactly to Gk. Σῆρες ‘silk men’; see the next paragraph). Whatever the exact details of this entanglement, there can be no doubt about a starting-point in East Asia, as far as its identity as a culture-word is concerned.

<sup>6</sup> “Written Mongolian” (or Literary Mongolian) is the scholarly term for an independent Mongolian language variety, attested in the old Mongolian script and different from both Classical Mongolian and Middle Mongolian. Although documented from the Middle Mongolian period, it represents an earlier linguistic stage.

### 3 The Silk Men

Neither is it at all disputed (e.g. Frisk 1960-66: 697) that Gk. σηρικός ‘silken, of silk’, substantivated σηρικόν ‘silk’ and transferred into Latin as *sēricus*, *sēricum* respectively, is formed from the same root as Eng. *silk*. It referred however not only directly to the raw silk and silk fabric, but also to its oriental origins. The Σῆρες<sup>7</sup> ‘silk men’ were initially of unknown ethnicity to the Greeks and Romans because they bought the silk via middlemen who transported it to the west from trade stations in Central Asia. Consequently, the “land of the *Sērēs*” acquired legendary associations until geographical knowledge of the Orient increased in the first centuries AD (Genaust 1996: 578-579; Thau-Knudsen 2000). At least Ptolemy knew in his *Geographia* from 150 AD that their land stretched beyond the *Imauni*, i.e. the Pamir mountains in Central Asia. The Chinese sequence *sī rén*, or whatever exact form was the starting-point, must have been transferred via Central Asiatic peoples to Europe and European languages; it cannot have been borrowed directly by European languages. The question then remains: What kind of Central Asians?

There is clear evidence that not only Europeans used the ‘silk’-term to denote the peoples of Central Asia, and that the term was not only used about traders, but about entire populations, again matching the translation ‘silk people’. First, on the famous Tonyukuk monument from 720 AD, found on the right bank of the upper course of the Tuul (Tola) river in Mongolia, the third and the fourth line of the Western side reads “*türk sir bodun yerintä / bod qalmadı*” ‘in the land of the Turkic *sir* people, no group of people were left (any longer)’ (Ramstedt, Granö & Aalto 1958: 30-31), and the term is mentioned again later in the inscription. Furthermore, the very ethnonym *Turk* in all probability derives from a translation of this name, cf. that similar words meaning ‘silk’ are found in all three Altaic branches: OTu. *torkū* ‘silk fabric’, MMong. *turge*, *torkan*, Written Mong. *torgan* ‘silk, satin’, Solon Evenki *tōrga* ‘silk’<sup>8</sup>.

<sup>7</sup> Transliterated *Sēres*, but the Latin form is *Sērēs*.

<sup>8</sup> Cf. also Fi. *turkku* ‘fur’ and ORu. *\*torgū* ‘marketplace, square’, borrowed into ON *torg* id. I do not necessarily embrace Wang’s (1993) idea that *Turk* and *silk* are ultimately etymologically identical, but in the light of a series of interesting of loanword proposals showing an alleged loan correspondence *\*sVIC-* vs. *\*tVrC-* the idea seems at least worth pursuing.

While it seems clear to everyone that the root-element *sil-* of the word *silk* and its congeners must be identical to the *ser-* in Gr. σηρικόν and the *Sērēs*, and that the source of this element is Chinese, it has not yet been cleared out how an *-ilk-* form might have arisen next to the *-ērik-* form, and why the variant with *-l-* is found exactly in some Germanic languages, in Baltic and in East Slavic (leaving aside the aberrant South Slavic forms *svila*)<sup>9</sup>. In fact, etymological handbooks in general seem to avoid explaining the variation. Vasmer (1953-58, III: 387) surmises that the Germanic forms are borrowed directly from Latin, which, however, does not explain the lambdacism, and that the Old Russian form, because of its initial consonant *š-* which constitutes yet another difference from the Mediterranean form, must rather have come from some unknown Eastern source<sup>10</sup>. While it is very plausible that the first intermediate stations were Altaic (Turkic, Mongolian, Tungusic) and perhaps Tocharian (Wang 1993, cf. also Hilmarsson 1984 on Tocharian *śorkām* ‘string’), pointing to the nomadic peoples of North-Western China and the Tarim region, no language of these groups preserves a form with *-l-* and the meaning ‘silk fabric’. What we are searching for is therefore a language at the Western End of the Silk Road that can have transformed *-r-* into *-l-* and which, at least linguistically, played a crucial role in the transmission of silk from east to west.

#### 4 A Northeast European isogloss

The geographical distribution of the *silk*-word in Europe provides some clues to its further migrations. As can be seen from the above list, it occurs in Germanic languages (from where it has been borrowed into North-Western Balto-Fennic and Western Saami), in Baltic languages, and in East Slavic (borrowed into North-Eastern Balto-Fennic, Southern Balto-Fennic and Eastern Saami). In Western Europe, the successor of Latin *saeta* ‘stiff hair, bristle’ has become the dominant term, however

<sup>9</sup> *Svila*, found in all South Slavic languages (Slovenian, Serbo-Croatian, Macedonian and Bulgarian) is technically unrelated since it formally seems to derive from a preverb *\*sŭ-* and the feminine of a participial form *\*vila* of the verb *\*viti* ‘to wind, to roll, to twist, to bend’, cf. e.g. Slov. *zvila* ‘bent’ (fem.), but this is likely to be due to folk-etymology (see Bezljaj 1995: 351 with references).

<sup>10</sup> Russian *š-* is regular from *s-* in Nordic loanwords; thus, it seems more likely that the East Slavic terms have been borrowed from Old Norse (Miller 2012: 66-67).

from the expression *saeta sērica* ‘silk thread’ (lit. ‘silk hair’)<sup>11</sup>. Thus, MHG *Seide*, Du. *zijde*, Sw. *siden*, Latv. *zīds*, Est. *siid* (in the two latter via MLG, from Latv. further into Liv. as *zīd*’) and terms in the Romance languages (Fr. *soie*, Sp. and Port. *sedá*, It. *seta*) are all ultimately, the Romance forms directly, from Lat. *saeta* and thus unrelated to the word *silk* despite superficial similarities. In other words, the *l*-form seems to be confined to Northern and Eastern Europe – and the full form involving both *-l-* and *-k-* is certainly only found there<sup>12</sup>.

During the Viking Age, silk was brought to Northern Europe by the Varangians, Nordic merchants who traveled through Kiev Rus and reached Constantinople where they traded with local merchants. One prominent example from the Icelandic sagas is the account of the Norwegian king Sigurd the Crusader’s visit to Miklagard (Byzantium) 1110, recorded in Snorri Sturluson’s *Heimskringla* (III, 238, cf. Blöndal 2007: 136), and from Nestor’s Chronicle we know how Russian envoys were gifted with impressive silken brocades during a visit to Byzantium 912 (Krag 2013). Given the geographical distribution of the *silk*-word and our knowledge, it is likely that the Varangians brought the word with them through Russia all the way to Scandinavia.

## 5 An Alanic sound-law

To sum up so far, although it is undisputed that the Chinese silk word must be the ultimate source of the several similar European terms for ‘silk’, the reason for the irregular variation between them has remained a mystery, and the exact languages that transferred them into Europe via Central Asia have resisted discovery. While inherited words

<sup>11</sup> Lat. *sērica* survives only via the Llat. form *sareca* in OFr. *sarge*, Fr. *serge* in the meaning ‘twill worsted’ or ‘twill silk’, and, via MLat. *sarcia* and OE *s(i)erc*, *sycr*, *s(i)erce* ‘shirt’ in ON *serkr* (Da. *særk*, Sw. *särk*, Nw. *serk*) ‘undergarment of silk or flax canvas’. Today, the Danish word has either historical or pejorative connotations, depending on context; in the latter case, the meaning is less specific and can refer to any kind of loose garment. The directions of transmission involving ME *serk*, Mod.Eng.dial. *sark*, OCS *sraky*, *sraka*, *sračica* ‘garment’, Ru. *soročka* ‘shirt’, id., Lith. *šárkas* ‘garment’, and Balto-Fennic forms like Fi. *sarkki*, Est. *särk*, Liv. *serk*, pl. *sērkid* ‘shirt’, is not entirely clear (cf. de Vries 1962: 471 with discussion and references).

<sup>12</sup> The exact relation of Pashto *sālū* ‘silk veil’ and Mod.Pers. *sirah* to *silk* as well as to each other is uncertain (Cheung 2002: 254, Vasmer 387), but cf. below on other Iranian terms.

generally obey to the regularity of sound-laws, culture-words, because of their tumultuous history as terms for migrating products and inventions, typically occur in numerous shapes whose internal relations are hard and sometimes impossible to disembroil. In the case of ‘silk’, the geographical distribution of the varieties may hold the key to a solution. We have seen how the *-l*-variants are found primarily in Northern and Eastern Europe, notably in Old English, Old High German, Old Norse, Baltic and Old Russian, while the *r*-variants have basically spread from Greek. A scenario where this variant of the silk-word migrated northwards from Byzantium through present-day Russia and Ukraine to Central and Northern Europe is compatible with the historical records. Since Chinese (and Manchu, Mongolian etc.) word-forms start out with *-r-*, it seems reasonable to assume that the Greek form of the word would be the more archaic one. What we search for is then a language near the Western end of the Silk Road that would regularly have transformed *-r-* into *-l-* before the 7<sup>th</sup> century AD<sup>13</sup> (but possibly much earlier), and which would be sufficiently culturally important as to provide or disseminate important culture-words.

Crucially, the regular development of older (Proto-Iranian) *\*-ri-* and *\*-r̥i-* in the subgroup of Scytho-Sarmatian languages within Iranian, comprising Scythian, Sarmatian, Alanic, Iassic and Ossetian, is exactly *-l-* (Bielmeier 1989: 241). This can be seen, first and foremost, in the name of the Alans themselves, Gk. Ἀλανοί, since it derives from the same (Indo-)Iranian self-designation *\*ariāna-* that has yielded the very name of *Iran* and the infamous politically misused term *Aryan* (originally, and among linguists and historians still, denoting Indo-Iranian, but not other Indo-European peoples). We know that this name does not derive from a language that turned *r* to *l* in general because a subgroup of the Alans were called Πωχολανοί which must reflect PIr. *\*rauxs-alana-* ‘shining Alans’ (where *\*rauxs-* in itself comes from PIE *\*leuk-s-* with regular Iranian development of *\*l- > r*), and, more indirectly, by the amalgamated name Ἀλανορσοί, mentioned by Strabo, where the second member is the name of the other subgroup, the Ἀορσοί whose name derives from PIr. *\*aruša-* ‘white’ (Mod.Oss. *urs* / *ors*; Cheung 2002: 7). At the same time, it known to have been an endonym, i.e. the Alans’ name for themselves. A widespread alternative name, originally an exonym but later adopted as an endonym, was Ἀσίοι, reflected in the

<sup>13</sup> When silk first occurs on the British Isles, simultaneous with the Old English velar umlaut of *\*siluc* to *seoluc*, cf. Miller 2012: 67.

name of Iassians, the Iazyges and the Ossetes, with no certain etymology.

The PIr. form *\*āryan-* is also indirectly attested as an Alanic loanword in the NE Caucasian Nakh languages: Chechen *ēla* ‘prince (of a principality), chieftain’, Ingush *āla* id. and Tsova-Tush (Batsbi) *ālē* ‘lord, gentleman’ can be reconstructed with a Nakh protoform *\*alä* that, in all likelihood, was borrowed from Alanic *\*ālan-* < PIr. *\*ārjan-*, cf. Oss. *Allon*, a mythological tribal name (Bielmeier 1989: 243; Thordarson 2009). Furthermore, the term occurs with the typical Iranian suffix *\*-ka-* in Oss. *Alægataē* (a Nart tribe) < PIr. *\*ārjaka-* (Thordarson 1989: 478).<sup>14</sup>

Another old attestation of the sound-law occurs in Herodotus’s Scythian name Κολάξαϊς since this most likely comes from PIr. *\*x<sup>w</sup>arja-kšajā* ‘sun-king’. Alkman’s ἵππος Κολαξαῖος shows that the development had taken place at least in the 7<sup>th</sup> century BC (Hinge 2005).

One further old example of the development *\*-ri-*, *\*-rj-* > Alanic *\*-l-* comes from the much later bulk of Alanic (Iassic) loanwords in Hungarian, transferred via the Iassic settlements in the 13<sup>th</sup> century. Hung. *zöld*, *zeld* means ‘green, unripe’ and is borrowed from Alanic *\*zalda-* < PIr. *\*zarita-* (corresponding to Ved. *hárīta-* and ultimately the same PIE formation as Eng. *gold*, only with different ablaut). Most examples of the sound-law are from Modern Ossetian, e.g. the preverb *fæl-* < *\*pari-*; Iron dialect *mælin*, Digor *mæluŋ* ‘to die’ < *\*mrja-* (the same IE root as in Lat. *mortuus* ‘dead’, and, as a loanword from French, Eng. *mortal*); Digor *zældæ* ‘young grass; grass; turf’ < *\*zarita* ‘yellow’ (~ Av. *zairita-*, Ved. *hárīta-* ‘yellowish’ and, with a different ablaut grade, Eng. *gold*), and *næl* ‘male’ (Digor originally *nælaē*) < *\*narja-* (~ Av. *nairiia* ‘male, virile’).

This means that if a foreign word containing the sequence *\*-ri-* (or *\*-rj-*) was borrowed into Alanic or another stage of Scytho-Sarmatian early enough, it would have yielded *\*-l-* in Alanic itself by regular sound development. It would have to have happened in antiquity already since the Alanic self-designation with *-l-* is mentioned by Strabo and Ptolemy. According to Pliny the Elder (*The Natural History*, 6,49), both the Jaxartes river and the Tanais (present-day Syr Darya and Don respectively) were called *Silis* by the Scythians, suggesting a

<sup>14</sup> The Sarmatian names Ἰρβιδος and Ἰργανος, attested in inscriptions from the Eastern part of the area North of the Black Sea, are perhaps also developments of *\*ārya-*, but they occur in the Western (“Scythian”) part as the variants Ἡλμάνου, Ἡλμάνος, which may reflect a dialect continuum where the change into *\*l* was not yet completed in the Sarmatian area (Hinge 2005).

development of \*-r- to “Scythian” -l-. Even though it is far from certain that the name of Syr Darya has any etymological affiliation to Silis, at least it shows that Scythian already possessed an l-sound.

## 6 The Iranian suffix \*-(i/a/ā)ka-

Furthermore, while the Iranian suffix \*-ika- formally corresponds to Greek \*-iko-, it became productive in the formation of nouns in Iranian languages<sup>15</sup>, most often however in the form \*-āka, \*-aka (originally formed to stems in \*-ā- and \*-a-) or plain \*-ka-, and was often added to words that would otherwise be or end up as monosyllabic. Thus, if a word \*sir or \*ser was borrowed into Alanic before the operation of the sound-law, or while it was still productive, it is very likely to quickly have formed part of a new derivative \*sirika- which would then develop into \*silka-. It would have stayed as \*silk(a)- until it was finally picked up by Varangians and other traders from the North in Byzantium.

Ciancaglini (2012a: 27-28 and 2012b: 95) notes about the Old Persian use of the suffix: “It seems that it occurs especially in toponyms and ethnonyms designating non-Iranian peoples, or peoples geographically distant or little known to the Persians”; among the examples she mentions *Karka* adj. ‘Carian’, *Karkā* ‘Caria’ where the suffix has been added to the original stem, but also e.g. *Katpatuka-* m. ‘Cappadocian’ where the final syllable in the foreign name seems to have been interpreted as the Old Persian suffix. It is at least conceivable that the word could spread via Old Persian areas with the meaning ‘(the somewhat remote) silk people’ even if it was formed in Scythian-Alanic itself directly to \*Sil-i-ka- ‘people of Silis, i.e. the area of Tanais and Jaxartes’. Based on what little we know about Scytho-Sarmatian and Alanic, we would at least expect a word \*silika ‘silk’ to be homonymous to an ethnonym (which, though hypothetical, is a possible productive formation). From the Black-Sea inscriptions we now know \*-(a)ka- in Sarmatian names (Hinge 2005), at least in the variants Φλιμάνακος, Φλειμνάγου, Φλειμναγος, Λιμνάκου and Λίμνακος, representing Scytho-Sarmatian \*(F)liyamanak/gos, formed with this suffix to the Iranian appellative \*friia-manah- ‘liebgeistig; having a loving mind’<sup>16</sup>,

<sup>15</sup> Except in Old Avestan where it is practically absent due to its low sociolinguistic connotation (Ciancaglini 2012a).

<sup>16</sup> Corresponding to Av. *friia-* ‘dear’, Ved. *priyā-* id., Eng. *free* + Av. *manah-* ‘soul, spirit’, Ved. *mānas-*, Gk. μένος.



which, in turn, is attested in Sarmatian Λείμανος and Mod.Oss. *lymæn* ‘friend’. It may even be attested via the Scythian name of the Amazons mentioned by Herotodus, Οιορπατα or Αιορπατα (with alternating diacritica) if Hinge is correct that ΟΙΟΠΠΑΤΑ is simply a corruption of \*ΟΙΟΠΓΑΤΑ < \*aiuauaragāta ‘one-breasted’ (cf. Av. *aēuua-* ‘1’ and *var-ah* ‘breast’).

Another possibility is that a Greek loanword reflecting σπηκόν was already around in Roman-age Alanic and that the Alanic word was simply formed on the basis of a Greek model where the Greek suffix *-ikó-* was interpreted as the Iranian *-(i)ka-* (cf. the parallel in OP *katpatuka-* ‘Cappadocian’ above). Even if we interpret Alanic \**silka-* as a Greek loanword lock, stock and barrel, it could still have been borrowed early enough for the sound-law to operate<sup>17</sup>. It is also possible that, after the completion of the sound-law, new sequences of \**-ri-* and \**-rj-* were so slow to reappear that *-l-* simply was the realization of a phonemic sequence /ri/, /rj/ for some time – meaning that if the language absorbed a foreign word with \**-ri-* or \**-rj-* it would automatically have this sequence substituted with phonetic [l]. However, we know that this is at least no longer the case in Byzantine 13<sup>th</sup> century Alanic where new cases of *-ri-* emerged from an old sequence *-raCj-*, cf. *zærin* or *zirin* < ζιρην > ‘yellow, golden’ (Mod.Oss. *zærin*) < Plr. \**zaranja-* (Cheung 2002: 104, 254; Engberg & Lubotsky 2003: 43).

A third possibility is that the culture-word indeed entered Alanic from some other language than Greek, but a language which nonetheless already contained an unrelated suffix with \**-k-*, cf. that such forms are already known in words for ‘silk’ or ‘thread’ from both Mongolic and Tocharian. Turkic in fact possesses a widespread suffix \**-lek* (\**-lak* after back vowels), which is already reconstructed in a textile term (Lubotsky & Starostin 2003: 261), namely PT \**köjñe-lek* ‘shirt’ (attested both in Modern Turkic languages and in loanwords in both Tocharian and Mongolian, and occurring in Altaic cognates with the meaning ‘silk’, ‘string’ and ‘thread’). A Turkic form \**sirlek* is likely to have turned into Alanic \**silka* via \**sirlika*. Also in this case the word could either have been interpreted by the Alanic borrowers as a completely foreign element and assigned a local phonetic rendering

<sup>17</sup> A loan from Byzantine Greek could account for the *-i*-vocalism (Miller 2012: 67), but it is uncertain if this is chronologically compatible with the consonantal development of \**ri* to *l*. At least it would rule out a connection with the province of Silis because in this name the development of \**r* to *l* would have taken place many centuries earlier.

(\**silka*), or it could have been reanalyzed by the speakers as a local formation with an inherited suffix (\**sirl-ika* > \**silka*), which is a perfectly normal development in loanword processes.

In any of these cases, the special Alanic sound-law \*-ri- > \*-l- plays an instrumental role. Despite the many uncertainties and possible subscenarios, it seems likely that the reason for the existence of the -*ilk*-form and its confinement to Northern and Eastern Europe is that it spread here via some stage of Alanic. Whether the derivative is formed with an Iranian suffix or has been borrowed with the suffix into Alanic from Turkic is less certain, as is its exact relationship with the Greek derivative, but it is important to note that the cultural and historical implications of all three scenarios are basically the same.

## 7 Golden silk – a reflex in Modern Ossetian?

An Alanic \**silka*- is not directly attested, not even in Iassic or Modern Ossetian. This is no serious problem since the loss and replacement of lexical items in a given language is a phenomenon that takes place in all languages over time. However, it is interesting that the Modern Ossetian word for ‘silk’, while not corresponding regularly with the Alanic reconstruction, is at least superficially similar and has no known alternative source. It occurs both as *zældag*, *zældagæ* ‘silk’ (Abaev IV) and without the \*-āka suffix as *zæly*, Digor *izæly* ‘silk; of silk, silken; silk scarf’ (Cheung 2002: 254). According to Abaev, Digor *i*- here is simply the definite article although prothetic *i*- elsewhere can have other sources (a Proto-Iranian syllable \**ja*- or the prepositions \**ui*- and \**abi*-).

Cases of Ossetic *z*- normally come from PIr. \**z*- (e.g. *zærdæ* ‘heart’ < PIr. \**zarita*-, *zad* ‘born’ < PIr. \**zāta*-; *zæronð* ‘old, old man’ < PIr. \**zaranta*-). On the other hand, there are quite a number of Ossetic words of unexplained origin beginning with *z*-. e.g. *zaz* ‘yew, Taxus’ and *zyng*, *zing* ‘burning, glowing hot coal; fire’. Some of these are in fact typical culture-words that seem to have had an original *s*-, e.g. *zæppaz* ‘elevated or submerged grave-chamber of stone’, if this is akin to Mod.Pers. *sabad*, Arm. *sap* ‘at’ ‘basket’ (cf. Cheung 2002: 254).

However, we would have expected Alanic \**silka*- to develop regularly in Ossetic since Ossetic derives directly from Alanic. One alternative possibility in the particular case of ‘silk’ is that *zæly*, *zældag(æ)* can have been influenced by the precursor of Oss. *zældæ* ‘young grass; grass; turf’ (Cheung 2002: 253). This implies that the meaning must still have been

‘golden’ at the time of reshaping, and we know that as late as Iassic this was still the case because the word has been borrowed into Hungarian as *zöld* with that meaning. ‘Golden’ would then either have referred to a particular golden type of silk garment, for example Byzantine embroidery, or to the golden Muga silk type from Assam, which reached the West alongside Chinese silk, or simply to the value or the glistening appearance of silk in general<sup>18</sup>. Such a contamination or folk-etymological reshaping could clearly have happened at any time in the history of Alanic from the time of its absorption of the supposed silk-term *\*sirika-* till the semantic narrowing of *zældæ* from ‘golden’ into ‘yellow grass’, since apart from the supposed Alanic loanword in surrounding languages, only the Modern Ossetian output *zældag(æ)*, *(i)zæly* is known.

## 8 Concluding remarks

Whether Oss. *zæly*, *zældag* indirectly continue Alanic *\*silka-* or not, a hypothetical Alanic form remains the most plausible mediator between Central Asia and ON *silki*, not least in view of the East Slavic attestations. Historically, this points to medieval trade routes on the Northern shores of the Black Sea. While the Alanic shape of the word seems to have reached Western Europe from at least in the 7<sup>th</sup> century AD (judging by the phonological shape of the OE forms) it is possible that the Alans imported it northwards in several waves, and it might not have reached Old Norse until the 9<sup>th</sup> century around the birth of Kievan Rus’ and the first attestation of the word in Germanic literary sources.

As can be expected, the export of silk terminology from China westwards comprises not only the word *silk* but a range of terms from this domain whose exact meaning are not always entirely clear (see., e.g. Ching 2011 on Tocharian *kaum* and other terms, and Lubotsky and Starostin 2003: 261 on the cognates of PT *\*köjñe-lek*). Tracking the exact routes of such Wanderwörter and their chronology requires minute phonological and morphological analyses compared with facts from cultural history.

<sup>18</sup> In the tenth-century Old English medical work now known as Bald’s *Leechbook*, jaundice is said to cause the body of the patient to ‘turn yellow like good yellow silk’ (*āgeolwāþ swā gōð geolo seoluc*; Biggam 2006: 3).

### Postscript

This article arose from discussions I had with the Danish balkanist Erik Thau-Knudsen who contributed with an essay on the history of silk terminology to the Danish National Encyclopedia (Thau-Knudsen 2000) of which I was the editor of etymologies at the time. We did not reach a satisfactory solution, although Thau-Knudsen hinted at another Iranian language, Parthian, as the provider. Jens Elmegård Rasmussen (p.c.) gave us the tip that Ossetian, with its frequent development of original *\*r* to *l*, might have played a role. On May 20, 2000, I presented the idea at a symposium at the University of Copenhagen (*Komparativ Sprogforskning på Vej*) that Alanic might have been the provider not only of the *l*-variants, but even the derivative itself, seeing that a) *\*(V)ka-* is a frequent nominal suffix in Iranian, b) the development of *\*r* to *l* was in fact older than Ossetian proper, and c) this sound-law does normally not work for Iranian *\*r* alone, but involves a following *\*-i-* that we also find in the Greek 'silk'-word. I reconstructed the same form *\*silika-* as Thau-Knudsen did, only for Alanic, and was now able to account for both morphological and phonological developments on the basis of our knowledge of Iranian in general and Alanic in particular. Only at the very end of editing the present article, I discovered that C. Gary Miller in his recent book *External Influences of English: From its Beginnings to the Renaissance* presents an almost identical solution (Miller 2012: 66-67), in fact with additional details; for example, he notes that *š* is regular in East Slavic borrowings from Old Norse. He suggests, too, that Oss. *zæly* derives from *\*silika-* (which he reconstructs with long *\*ī*)<sup>19</sup>, however without explaining the irregularities or mentioning the important variant *zældag*. It appears to be time for etymological dictionaries to revise their entries on *silk*.

### References

- Abaev, Vladimir I., 1958-1989. *Istoriko-etymologičeskij slovar' osetinkogo jezika*: I-IV. Moscow / Leningrad.
- Bezljaj, France, 1995: *Etimološki slovar slovenskega jezika*, vol. 3 (P-S). Ljubljana: Izdala Slovenska Akademija Znanosti in Umetnosti & al. (edited by Marko Snoj & Metka Furlan).

<sup>19</sup> On the basis of Byzantine Greek /siri'kon/; its length is indirectly preserved in Old Irish *síric* (via Vulgar Latin *\*sīrica*).

- Bielmeier, Roland, 1989: "Sarmatisch, Alanisch, Iassisch und Altossetisch". – Rüdiger Schmitt (ed.): *Compendium Linguarum Iranicarum*. Wiesbaden: Reichert. Pp. 236-245.
- Biggam, Carole, 2006: "Old English colour lexemes used of textiles in Anglo-Saxon England". – Graham D. Caie, Carole Hough and Irene Wotherspoon (eds.): *The Power of Words: Essays in Lexicography, Lexicology and Semantics In Honour of Christian J. Kay*. Amsterdam: Rodopi. Pp. 1-21.
- Blöndal, Sigfús, 2007: *The Varangians of Byzantium*. Cambridge: Cambridge University Press.
- Cheung, Johnny, 2002: *Studies in the Historical Development of the Ossetic Vocalism*. Wiesbaden: Ludwig Reichert Verlag.
- Ching Chao-Jung, 2011: "Silk in Ancient Kucha: On the Tocharian B word *kaum*\* found in the documents of the Tang period. *Tocharian and Indo-European Studies*. 63-82.
- Ciancaglini, Claudia, 2012a: "Il suffisso Indo-Ir. *\*-ka* nelle lingue iraniche antiche". – *Archivio Glottologico Italiano* XCVII, fascicle I: 3-33.
- Ciancaglini, Claudia, 2012b: "Outcomes of the Indo-Iranian suffix *\*-ka* in Old Persian and Avestan". – Gian Pietro Basello & Adriano V. Rossi (eds.): *Dariosh Studies II, Persepolis and its Settlements: Territorial system and archaeology in the Achaemenid state*. Napoli: Università degli studi di Napoli "Orientale" (= *Associazione Internazionale di Studi sul Mediterraneo e l'Oriente, Series Minor* LXXVIII). Pp. 91-100.
- Engberg, Sysse & Alexander Lubotsky, 2003: "Alanic marginal notes in a Byzantine manuscript: a preliminary report. – *Nartamongæ: The Journal of Alano-Ossetic Studies* II, 1-2: 41-46.
- Falk, H.S. & Alf Torp, 1960: *Norwegisch-Dänisches etymologisches Wörterbuch*. Oslo / Bergen: Universitetsforlaget.
- Fraenkel, Ernst, 1962-1965: *Litauisches etymologisches Wörterbuch* I-II. Heidelberg: Carl Winter.
- Frisk, Hjalmar, 1960-1966: *Griechisches etymologisches Wörterbuch* I-III. Heidelberg: Carl Winter.
- Genauast, Helmut, 1976 (3<sup>rd</sup> ed. 1996): *Etymologisches Wörterbuch der botanischen Pflanzennamen*. Basel / Boston / Berlin: Birkhäuser Verlag.
- Hilmarsson, Jörundur, 1984: "East Tocharian *śorkäm* 'peg' or 'string'? – *Die Sprache* 30: 16-28.
- Hinge, George, 2005 [2006]: "Herodot zur skythischen Sprache: Arimaspen, Amazonen und die Entdeckung des schwarzen Meeres". – *Glotta* 81: 86-115.
- Krag, Anne Hedeager, 2013: "Silkens vej til Norden". *SFINX* 2013,1: 24-29.
- Lubotsky, Alexander & Sergei Starostin, 2003: "Chinese and Turkic loanwords in Tocharian". – Brigitte L.M. Bauer & Georges-Jean Pinault (eds.): *Language in Time and Space. A Festschrift for Werner Winter on the occasion of his 80<sup>th</sup> birthday*. Berlin / New York: Mouton de Gruyter. Pp. 257-269.
- Miller, C. Gary, 2012: *External Influences of English: From its Beginnings to the Renaissance*. Oxford: Oxford University Press.
- ODEE = *Oxford Dictionary of English Etymology*, edited by C.T. Onions with the assistance of G.W.S. Friedrichsen and R.W. Burchfield. Oxford: Oxford University Press, 1966.

- Plöger, Angela, 1973: *Die russischen Lehnwörter der finnischen Schriftsprache*. Wiesbaden: Harrassowitz.
- Ramstedt, G.J., J.G. Granö & Pentti Aalto, 1958: "Materialien zu den alttürkischen Inschriften der Mongolei". – *Journal de la Société Finno-Ougrienne* 60: 3-91. Part I. "Die Grabinschrift von Tonyukuk".
- Schrader, Otto, 1904-1905. "Über Bezeichnungen der Heiratsverwandtschaft bei den indogermanischen Völkern" (11-36). Part "Slavische oder durch Slaven vermittelte Lehnwörter im älteren Deutsch" (29-36). *Indogermanische Forschungen* 17.
- SKES = 1958-1981: Yrjö H. Toivonen, Aulis J. Joki, Reino Peltola, Erkki Itkonen, Satu Tanner & Marita Cronstedt: *Suomen kielen etymologinen sanakirja* I-VII. Helsinki 1958-1981 (= *Lexica Societatis Fenno-Ugricae* XII): Suomalais-Ugrilainen Seura.
- Thau-Knudsen, Erik, 2000: "Silke – ordets historie". Etymological essay attached to the entry "silke" in *Den Store Danske Encyklopædi*, vol. 17. Copenhagen: Danmarks Nationalleksikon.
- Thordarson, Fridrik, 1989: "Ossetic". – Rüdiger Schmitt (ed.): *Compendium Linguarum Iranicarum*. Wiesbaden: Reichert. Pp. 456-479.
- Thordarson, Fridrik, 2009: *Ossetic grammatical studies* [= *Veröffentlichungen zur Iranistik* 48]. Vienna.
- Vasmer, Max, 1953-1958: *Russisches etymologisches Wörterbuch* I-III. Heidelberg: Carl Winter.
- de Vries, Jan, 1962: *Altnordisches etymologisches Wörterbuch*. Leiden: E.J. Brill.
- Wang, Penglin, 1993: "On the Etymology of English *silk*: A Case Study of IE and Altaic Contact". *Central Asiatic Journal* 37: 3-4. Pp. 225-248.



# English *mink*, Finnish *portimo* ‘ermine’, and Baltic Fur Trade from Antiquity to the Hanse

## Abstract

Eng. *mink* and its Germanic cognates are loanwords from Baltic where today in Lithuanian *menkė* means ‘cod’, while related Slavic forms denote the ‘burbot’, a closely related species of freshwater fish. This species shared its name with the weasel in both Latin (*mūstēla*) and Greek (γαλέη). In Baltic it is itself a borrowing from Proto-Mari *men* ‘burbot’ which is a Uralic name (cf. Hung. *mëny-hal* ‘burbot’ ~ *mënyét* ‘weasel’). This explanation finds support in Kalima’s (1936) etymology, deriving Lith. *šėškas* ‘polecat’ from Mari *šäška*, *šaške* ‘weasel’. The spread of these terms are likely to be connected to fur trade around the Baltic. It is further suggested that Fi. *portimo*, dial. *porttimo* ‘stoat, ermine’ is a derivative of a medieval loan from MDu. *furet*, Late ME *forette*, ultimately from Lat. *fūrō* ‘polecat’, brought to the North Baltic via the Hanseatic trade.

## 1 Introduction

Strikingly many Northern European terms for the group of predatory mammals referred to in biology as *mustelids* (weasel, stoat/ermine, polecat, ferret, mink, marten, otter, sable, wolverine, badger etc.) have no established etymologies, e.g.: Eng. *stoat* ‘ermine’; Eng., LG *mink*, Sw. *mänk*, *menk*, originally ‘European mink’; Fi. *kärppä* ‘stoat, ermine’; Lith. *šėškas*, Latv. *sesks* ‘polecat, ferret’; Fi. *portimo*, dial. *porttimo* ‘stoat, ermine’ and its cognates in Karelian, Lude and Veps; and Nw. *jerv* ‘wolverine’, just to mention a few. It is likely that this situation reflects early language contacts connected to fur trade along the coasts of the Baltic Sea from the Roman Ages onwards. This article seeks to cast new light on a few of terms and their etymologies. New proposals are presented



which have implications for our understanding of this part of European cultural history and its reflections in language.

## 2 Lithuanian *šėškas* ‘polecat, ferret’

Lith. *šėškas*, Latv. *sesks* ‘polecat, ferret’ were first compared to Ved. (YV) *kása-* and the Rigvedic hapax *kaśikā-* by Fick (1879: 165) who identified the correspondence Ved. *\*ī* ~ Balt. *\*-Ø-* as PIE *\*a*. These forms were later analyzed by numerous scholars (see Kalima 1936: 102–103) as a reduplicative formation of the root in *śikti* ‘to shit’, i.e. *\*še-ś(i)ka-s* with typical loss of *-i-*. The reconstruction of a PIE form *\*keḱ-* (*\*keḱ-*) on the basis of Vedic-Baltic comparison, despite the aberrant initial consonantism, has almost become a generally accepted etymology, mentioned as the only option in standard handbooks. However, the meaning of neither Vedic word is decidedly certain (Katz 2002: 303). Equally important, while numerous “kentum reflexes” occur in Baltic with regular counterparts in Slavic and Indo-Iranian, the reverse irregular correspondence Lith. *š-* ~ Indo-Iranian and/or Slavic *\*k-* is at best very rare. This is understandable since the centum-satem isogloss must have divided the Core Indo-European homeland<sup>1</sup> approximately in the middle, so that the languages spoken in the westernmost fringes of the Eastern parts absorbed more kentum forms than the other satem languages. One might ask whether examples of Baltic *š-* ~ Indo-Iranian *\*k-* other than the one discussed exist at all – and cases of Baltic *š-* ~ Slavic *\*k-* where *k-* does not precede a sonorant (an environment that would probably depalatalize it) are close to absent. Thus, if LCS *\*kərmb* ‘fodder’ (~ Lith. *šerti* ‘to feed’) really has *k-* because of the following sonorant, the only example seems to be Lith. *šeivà* ‘spool’ < *\*šaiya-* alongside Slavic *cěva* < *\*kaiya-* ‘id.’.

Thomsen (1890: 223) was the first to etymologize Veps *hähk* ‘otter’ and its Balto-Fennic cognates, Fi. dial. *häähkä*, Olonets Karelian *heähku*, Lude *heähkäine* ‘European mink’, as a loanword from the Baltic predecessor of Lith. *šėškas*.<sup>2</sup> Wichmann (Kalima (1936: 102) instead suggested that the Baltic word was the one that had been borrowed, and

<sup>1</sup> Like most scholars today, I use Core Indo-European about the protolanguage being left behind after the departure of Anatolian and Tocharian.

<sup>2</sup> Older Baltic loanwords with *\*š* show up with *h* in Balto-Fennic languages in line with the fact that Early Proto-Fennic *\*š* regularly yields *h*. \*Livonian *säsk* is a Latvian loanword.

that Balto-Fennic was the provider. He pointed to the fact that Mari (Cheremis) has a word *šäškä*, *šaškä*, meaning either ‘mink’ or ‘otter’ depending on the dialect – some dialects have *βüt-šäškä* (the 1<sup>st</sup> element being *βüt* ‘water’), meaning ‘otter’ only. This word corresponds nicely with the Balto-Fennic protoform *\*hähkä*, and Wichmann saw that the two must be cognates. His hypothesis was then that the borrowing had taken place at the Early Proto-Fennic stage, when the form would still have been *\*šäškä*.<sup>3</sup>

It has now become an established opinion (see the overview of research in van Pareren 2005, 2008) that Baltic exercised at least a modest influence upon Mordvin lexicon and toponymy. Thomsen (1890) and Kalima (1936) thought that even (Proto-)Mari had come into some contact with Baltic. Although Mägiste (1959), much to his own chagrin, was mostly negative, he did not exclude the possibility completely. Likewise, the inference that Baltic contributed to the Mari lexicon is implied by a handful of entries in UEW.

Volgaic languages<sup>4</sup> historically stretched further to the West, bordering areas inhabited by Baltic tribes. Conversely, Baltic languages reached much further East, not least reflected in the huge area covered by the Grand Duchy of Lithuania which extended almost all the way to Moscow.

Here I would like to propose the further possibility that lexical borrowings within a least one semantic field, animals coveted for their pelts, even took place in the opposite direction, from Mari to Baltic. Mari *šäškä*, *šaškä* ‘mink; otter’ is already known to have been borrowed into neighbouring Turkic languages (Kalima 1936: 102, Räsänen 1969: 105): Chuvash *šaškä* ‘European mink’, Bashkir *šäškä* ‘id.’, (dial.) *šaška* ‘marten’, Tatar *čäškä*, *čäškä* ‘a water animal’. It can still be a Fenno-Volgaic word with remnants in Balto-Fennic and Mari. But note the important

---

<sup>3</sup> Junttila (2012: 268) nonetheless seems to prefer Thomsen’s etymology, grouping it in his category A: “Relatively clear etymologies”.

<sup>4</sup> Most Uralicists now reject that Mari and Mordvin once formed a genetic subgroup of Uralic. I leave the question open, but in any case it can be used about a geographical group of Fenno-Ugric languages which must have been subject to some convergence (e.g. a tendency to dissimilate sequences of two nasals), especially since the ethnic groups are regarded to have constituted a historical entity. From Kievan Rus’ we know the names of three more westerly Volgaic groups, the Merya, the Muromians, and the Meshchorya. The latter survived at least into the 16<sup>th</sup> c., judging by Russian chronicles. Curiously, although the five Volgaic ethnonyms are likely to have at least four different etymologies, they all begin with *M-* (except that the non-native name of Mari is *Cheremis*).

fact that Mari has a similar word, Meadow Mari *šeške* ‘daughter-in-law; young lady’ (Paasonen/Siro 1948: 121), Hill Mari ‘daughter-in-law’ (Ramstedt 1902: 129); in some dialects the meaning is more specific, such as ‘the wife of one’s son’, ‘the wife of one’s younger brother’, ‘the wife of one’s husband’s younger brother’ and ‘the wife of one’s wife’s younger brother’ (Moisio & Saarinen 2008: 678). The use of words for ‘bride’, ‘daughter-in-law’, ‘sister-in-law’, or ‘(small) young female’ about weasels and other mustelids, historically connected to a symbolism for pristine, virginal and desired young women, is extremely widespread across the West Eurasian area (see e.g. Falk 1998: 93–94, Witczak 2004, Martirosyan 2010: 799–800, Olsen 2010: 16, fn. 17), e.g. Da. *brud* ‘bride; least weasel, *Mustela nivalis*’, PGmc. *\*marþu-* ‘marten’ ~ Lith. *marti* ‘bride’, Latvian *mārša* ‘sister-in-law’; It. *donnola* and Port. *doninha* ‘weasel, lit. little lady’, Sp. *comadreja* ‘weasel, little god-mother’, Basque *satandre* ‘weasel’, < *\*sagut-* ‘mouse’ + *andere* ‘lady’; Gr. γαλέη ‘weasel’ ~ γάλως ‘sister-in-law’, Mod.Gr. νυφίτσα ‘weasel; little bride’; Hung. *hölgy* ‘weasel, bride’, Ru. *kunica* ‘(little) marten; bride (in traditional wedding rituals)’, Arm. *hašn-owk* ‘weasel’ ~ *harsn-owk* ‘little bride’ and Turk. *gelin* ‘bride’, dim. *gelincik* ‘little bride, little young woman; weasel’ – just to mention a few. This is also the case in the cultures of Western Siberia from where furs were provided, e.g. the Komi (or Zyryans; Laakso 2005).

This makes it probable that Mari *šeške* ‘daughter-in-law; young lady’ is related to *šäškə, šaškə* ‘otter, mink’, at least indirectly. The former is likely to be formed as a diminutive of the word for ‘sister’, which today is KB *šəžar*, U *šüžər*, B *šužar*. Two protoforms of this word are reconstructed for Fenno-Volgaic, *\*sasare* (reconstructable also for the Fenno-Permian stage) and *\*sisare* or *\*sesare*; the former is regarded as Early Indo-Aryan or Indo-Iranian, the latter as Baltic, and the Mari word can come from either of them; the rounded vocalism in the first syllable is secondary from assimilation to the sibilant (UEW 752, 762). Mari has to velar diminutive suffixes; *-γə* and *-ka* ~ *-(i)kä*, going back to PU *\*-kV* and *\*-kkV* respectively. The former is realized as the allomorph *-kə* after a sibilant (Wichmann 1913–1918: 7–9, 11–13). It is therefore likely that both *šäškə, šaškə* ‘mink; otter’ and *šeške* ‘daughter-in-law; young lady; sister-in-law’ are diminutives of the ‘sister’-word, the former of *\*sasare* and the latter of *\*sesare*, at an early time when the vocalism in the Mari ‘sister’-word had not yet become rounded.

In the following, I will discuss another possible loanword from Mari in Baltic which ended up ultimately in English as the name of a North American species.

### 3 English *mink* and Lithuanian *menkė*, Sorbian *mjenk* ‘burbot’

English *mink*, attested as ME *menks*, *mynkes* from the 15<sup>th</sup>-16<sup>th</sup> c., and its cognates in Germanic, LG *mink* ‘otter’, Sw. *mänk*, *mink* ‘id.’, are of unknown origin. Today the English term denotes a North American animal, but it was originally used only of the European mink (*Mustela lutreola*, German *Nertz*). Etymological dictionaries typically treat it as an ancient culture-word connected to fur trade around the Baltic sea, without speculating further about its origins.

Lat. *müstēla* and Gk. γαλέη, γαλή both have a curious double meaning ‘weasel’ and ‘burbot’, a sweetwater fish species (cf. Schaffner 2006) with which it shares similarities both in visual characteristics and behavior. While the double meaning in Latin of course could have been taken over from Greek, connections between the two animals also show up in Eastern Europe. The Hungarian name for the burbot is *mëny-hal* (*hal* ‘fish’), lit. ‘weasel-fish’, cf. *mënyét* ‘weasel’, *mëny* ‘daughter-in-law; (OHung.) ‘bride’; and a West Slavic name for the ‘burbot’ is Sorb. *mjenk*, Cz. *mník*, whose Baltic cognate is Lith. *mėnkė*, Latv. *mēnca*, which was transferred to the salt-water cod, a word that obviously shares great superficial similarities with our *mink*.

It thus seems reasonable to establish the Germanic word for ‘mink’ as a loanword from a Baltic or a West Slavic language. Since the mink is often designated by its wet habitat (Lith. *audinė*, cf. Young 2001; Fi. *vesikko*, cf. *vesi* ‘water’; Da. *flodilder*, lit. ‘river polecat’) the recurrent element *\*min-* in Baltic river names comes to mind. But the line of transmission does not begin in Baltic or West Slavic. The formation mentioned above is also found in Slovenian *menək*, *menká* and derived from LCS *\*mьnbь*, reconstructed on the basis of Ru. *men’*, Ukr. *min’*, Slovak *mieň* and the rarer Czech simplex *meň*. The root *man-* in SCr. *manić*, curiously missing from Vasmer (1953-1958), also fits in (cf. *pas* ‘dog’ < *\*mьnbь*), while the Lechitic derivatives Pol. *miętus* (→ Belaru. *mjantúz*), Kashubian *mińtus*, as well as Rusyn *mn’uh*, Ru.dial. *ment’uk* (→ Moksha Mordvin *ment’uk*, Erzya Mordvin *\*mänt’uk*), are less well understood. These Balto-Slavic forms are normally regarded to constitute a word-family with PGmc. *\*muniwō* (> Eng. *minnow*) and Gr. μαίνη ‘small fish,

sprat'; a root *\*menH-* with the zero-grade of an i-stem form the basis of the Germanic, Common Slavic and Greek forms while the Baltic and West Slavic forms extended with a suffix *-k-* have full-grade. It is noteworthy, however, that the Germanic and Greek forms designate specifically small kinds of fish and could just as well have been formed independently from (Pre-)PGmc. *\*min-u-* 'small' and (Pre-)PGk. *\*mei-n-* 'id.' respectively. The Balto-Slavic forms are terms for a completely different kind of species, overlapping almost completely with Fenno-Volgaic forms: the aforementioned Hung. *mëny-hal* and Hill Mari (West Cheremis) *men*, *men-gol*, also 'burbot'.

At first glance it appears that Mari simply borrowed their terms for the fish from Russian *men'*, while the Hungarian word, first attested 1395, would have been taken over from Pannonian Slavic. However, there are several problems with this seemingly straightforward scenario. First, the Hungarian vocalism is wrong since the manifestation *-ë-* (which is not distinguished from *-e-* in the standard language) normally points to back vocalism in the stressed syllable of the lending language, but this Slavic word only had front vowels. Second, it would be an amazing coincidence, in the light of other European connections between terms for 'burbot' and 'weasel; young lady' if the Slavic loanword meaning 'burbot' just happened to have the same shape in Hungarian as the inherited word for 'daughter-in-law; weasel' (PU *\*miñä* 'daughter-in-law; young lady', UEW 276).

Moreover, the Hungarian and Mari words actually correspond to each other as from a common protoform, and Skolt Saami has *manij* '(big) whitefish, Coregonus (lavaretus)', *moanji*, *moanjigaž* 'id.' (Collinder 1977: 115; SKES 347-348). Since *\*šeškas* already constitutes a possible Mari loanword in Baltic, I find it plausible that the stem *\*men-* in the meaning 'burbot; weasel' was transferred to from Volgaic to the Baltic Sea region as well. Among other possibilities, it could have happened from Proto-Mari to Late Common Slavic and (East) Baltic. The suffixated Slavic forms could either be borrowings from Baltic or parallel formations. From burbot skin Fenno-Ugric forest peoples have traditionally produced clothing – coats, boots, and caps – justifying an analogy with terrestrial animals hunted for their pelt (Armstrong 1997). The burbot along with the sturgeon and sterlet were so important for the Khanty (Ostyaks) in the time of Russian expansions in the 1600s that a particularly bad fishing season could threaten the very existence of a tribe. Novgorod and Moscow exploited the fur resources of the North

for its foreign trade, and luxury furs from mustelids like ermine and sable were called the “gold of ancient Rus” (Platonov & Andreev 1922).

Whether the Baltic and Slavic forms are Indo-European or borrowed from Fenno-Volgaic, it seems clear that the Germanic term for the mink must derive from the Balto-Slavic fish-name. It is further conceivable that the term *mink* spread in Western Europe under the influence of Du. *minneken* ‘playful term for a female’ > Eng. *minikin* (attested from the 16<sup>th</sup> c.).

#### 4 Finnish *por(t)imo*

Leaving the question of Volgaic loans into Baltic, let us turn to the Finnish name of the ‘stoat, ermine’, *portimo*, dial. *porttimo*, and its cognates in Karelian, Lude and Veps. Previous proposals are unsatisfying: Koi-vulehto (1979) derives it from Fi. *porras*, gen. *portaan* ‘step; staircase etc.’ with reference to the animal’s habitat in human houses and farms; and Liukkonen (1999) imagines a rendering of a hypothetical Baltic †*sparteiva* ‘swift animal’, cf. *spartūs* ‘swift’<sup>5</sup>.

That *-ti-* fails to assibilate into *\*-si-* speaks for a Germanic rather than a Baltic origin, because most of the Baltic loanwords in Balto-Fennic were borrowed before the assibilation. Since the distribution is confined to the Northern branch of Balto-Fennic, a somewhat recent origin in Germanic is likely. BF *\*portti-* would be the expected substitution of Late ME *forette*, MDu. *foret* ‘ferret’, borrowed from OFr. *furet* < Vulg.Lat. *\*furittus*, lit. ‘little polecat’, derived from *fūrō* ‘polecat’. The name could have spread Northwards with the Hanseatic trade between the 13<sup>th</sup> and the 17<sup>th</sup> century, perhaps even from the end of the 12<sup>th</sup> century (See Bentlin 2008 for a new account on MLG loanwords in Finnish). Substituting *forett-* with *portt-* would be expected, cf. Fi. *per-jantai* ‘Friday’ alongside MLG *vrîdach*, Old Bavarian *pferintag* ‘id’. (Bentlin 2008: 156).<sup>6</sup>

The suffix *-imo* is not excessively common in Fennic but occurs in many place-names and a few nouns like *tuhkimo* ‘ashtray’ from *tuhka*

<sup>5</sup> Junttila (2012: 268) groups this proposal under his category B: “Dubious etymologies”, rather than in C: “Erroneous etymologies”.

<sup>6</sup> In fact one may wonder if Eng. *polecat* is not a folk-etymological reshaping of the word *\*pol’eka* that yielded e.g. North Saami *buoidaga* ‘weasel’, brought Southwestwards with the Hanseatic trade or perhaps earlier in connection with the fur trade.

‘ash’ (Koivulehto 1979; Hakulinen)<sup>7</sup>. However, *-mo* here may represent a Karelian dialectism for original *\*-ma*, cf. another name for a mustelid, Kar. *ohmo*, *ahmo* for Fi. *ahma* ‘wolverine’. In that case, the original distribution of the word would be confined to the Northeastern Balto-Fennic languages. This at first glance perhaps renders a Middle Dutch or Middle Low German origin less likely. One could however also argue the other way around – that a limited geographical distribution is an argument for a relatively late migration connected to urbanism – pre-historic and early historic vocabulary typically occur with a more scattered distribution. Likewise, the very limited morphological variation speaks for a historically late introduction of the word. When a non-native term has been present in a given area for a long historical period, we typically expect to encounter more kinds of domestic derivatives formed on the basis of this term, with different nominal suffixes.

## References

- Armstrong, Peter C.B., 1997: *Foreigners, Furs and Faith: Muscovy's Expansion into Western Siberia, 1581-1649*. MA thesis from Dalhousie University, Halifax, Nova Scotia.
- Bentlin, Mikko, 2008: *Niederdeutsch-finnische Sprachkontakte: Der lexikalische Einfluß auf die finnische Sprache während des Mittelalters und der frühen Neuzeit* [= Mémoires de la Société Finno-Ougrienne 256]. Helsinki: Suomalais-Ugrilainen Seura.
- Falk, Harry, 1993: “Der Zobel im Rigveda”. – Gerhard Meiser (ed.): *Indogermanica et Italica* [Fs. Helmut Rix = Innsbrucker Beiträge zur Sprachwissenschaft 72]. Innsbruck: IBS. Pp. 76-94.
- Fick, August, 1879: “Schwa indogermanicum”. – *Beiträge zur Kunde der indogermanischen Sprachen* 3: 157-165.
- Junttila, Santeri, 2012: “The prehistoric context of the oldest contacts between Baltic and Finnic languages”. – Riho Grünthal & Petri Kallio (eds.): *A Linguistic Map of Prehistoric Northern Europe* [= Mémoires de la Société Finno-Ougrienne 266]. Helsinki: Suomalais-Ugrilainen Seura. Pp. 261-296.
- Kalima, Jalo, 1936: *Itämerensuomalaisten kielten balttilaiset lainasanat*. Helsinki: Suomen Kirjallisuuden Seura.
- Koivulehto, Jorma, 1979: “Baltisches und Germanisches im Finnischen: die finn. Stämme auf *-rte-* und die finn. Sequenz *VrtV*”. – E.F. Schiefer (ed.): *Explanatio-nes et tractiones Fenno-Ugricae in honorem Hans Fromm*. München: Wilhelm Fink. Pp. 129-164.

<sup>7</sup> It is conceivable that the use of exactly this quite rare suffix in a late word for ‘ermine’ was furthered by association between *tuhkimo* and the only other word in Finnish beginning with *tuhk-*, namely *tuhkuri* ‘mink’.

- Kari, Liukkonen, Kari, 1999: *Baltisches im Finnischen* [= Mémoires de la Société Finno-Ougrienne 235]. Helsinki: Suomalais-Ugrilainen Seura.
- Katz, Joshua T., 2002: "How the Mole and Mongoose got their Names: Sanskrit *ākhú-* and *nakulá-*". – Joel P. Brereton & Stephanie W. Jamison (eds.): *Indic and Iranian Studies in Honor of Stanley Insler on his Sixty-fifth Birthday* (= *Journal of the American Oriental Society* 122, 2 [special issue]). Pp. 296-310.
- Laakso, Johanna, 2005: *Our otherness: Finno-Ugrian approaches to women's studies, or vice versa* [= *Finno-Ugrian Studies in Austria* 2]. Münster: LIT Verlag.
- Mägiste, Julius, 1959: "Gibt es im Tscheremissischen baltische Lehnwörter?". – *Ural-Altaische Jahrbücher* 31: 169-176.
- Martirosyan, Hrach, 2010: *Etymological Dictionary of the Armenian Inherited Lexicon*. Leiden / Boston: Brill.
- Moisio, Arto & Sirkka Saarinen, 2008: *Tscheremissisches Wörterbuch* [= *Lexica Societatis Fenno-Ugricae XXXII*]. Helsinki: Suomalais-Ugrilainen Seura.
- Olsen, Birgit Anette, 2010: *Derivation and composition: Two studies in Indo-European word formation* [= *Innsbrucker Beiträge zur Sprachwissenschaft* 136]. Innsbruck: IBS.
- Paasonen/Siro 1948 = H. Paasonens *Ost-tscheremissisches Wörterbuch*, edited and revised by Paavo Siro [= *Lexica Societatis Fenno-Ugricae XI*]. Helsinki: Suomalais-Ugrilainen Seura.
- van Pareren, Remco, 2005: *Baltische leenwoorden in het Mordwiens*. MA thesis from the University of Groningen.
- van Pareren Remco, 2008: "Die direkten baltischen Lehnwörter im Mordwinischen". – *Finnisch-Ugrische Mitteilungen* 30-31: 69-147.
- Platonov, F.S. & S.I. Andreyev, 1922: "Očerki novgorodskoi kolonizatsii severa". – *Očerki po istorii kolonizatsii severa*. Saint Petersburg: Gosudarstvennoe izdatel'stvo. Pp. 26-37.
- Ramstedt, Gustaf John, 1902: *Bergtscheremissische Sprachstudien* [= *Mémoires de la Société Finno-Ougrienne XVII*]. Helsinki: Suomalais-Ugrilainen Seura.
- Räsänen, Martti, 1969: *Versuch eines etymologischen Wörterbuchs der Türksprachen* [= *Lexica Societatis Fenno-Ugricae XVII*]. Helsinki: Suomalais-Ugrilainen Seura.
- Schaffner, Stefan, 2006: "Lateinisch *müstella*, *müstēla* 'Wiesel; Quappe' und der Wortbildungstyp vedisch *aśvatará-*". – *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 3: 1-50.
- Suhonen, Seppo, 1988: "Die baltischen Lehnwörter der finnisch-ugrischen Sprachen". – Denis Sinor (ed.): *The Uralic Languages: Description, History, and Foreign Influences*. Leiden / New York / Copenhagen / Cologne: E.J. Brill. Pp. 596-615.
- Thomsen, Vilhelm, 1890: *Beröringer mellem de finske og de baltiske (litavisk-lettiske) Sprog. En sproghistorisk Undersøgelse*. Copenhagen: Det Kongelige Videnskabskabernes Selskab.
- Vaba, Lembit, 1983: "Baltische Lehnwörter der Wolga-Sprachen im Lichte neuerer Forschungsergebnisse". – *Sovetskoie Finnougrovedenie* 19: 138-145.
- Vasmer, Max, 1953-1958: *Russisches etymologisches Wörterbuch*. Heidelberg: Winter.
- Wichmann, Yrjö, 1911: "Zur Geschichte der finnisch-ugrischen anlautenden Affrikaten bes. im Ungarischen und im Finnischen". – *Finnisch-Ugrische Forschungen* 11: 173-290.



- Wichmann, Yrjö, 1913-1918: "Beiträge zur tscheremissischen Wortbildungslehre". – *Journal de la Société Finno-Ougrienne* 30, 6: 1-42.
- Witczak, Krzysztof Tomasz, 2004: "Old Prussian *moazo* 'mother's sister', *mosuco* 'weasel' and related words". – *Baltistica* 39 (1): 131-139.
- Young, Steven, 2001: "East Baltic terms for 'wolverine' and 'mink'". – *Linguistica Baltica* 9: 163-167.

# **Latin and Slavic Loanwords in Hungarian: Exceptions to Helimski's Vowel-Harmony Adaptation Rule**

## **Abstract**

Eugen Helimski (1998) showed that the accented vowel of High German, Latin and Slavic words governs whether they acquire front-vowel or back-vowel harmony as loanwords into Hungarian. While the German material appears exceptionless, some words of Slavic and Latin provenance exhibit unexpected back-vowel harmony. This article shows that if a labial sound follows the originally accented vowel, front-vowel harmony is blocked. The rule applies without exception to both Slavic and Latin loanwords; it is thus an economical solution.

It follows that variation in Slavic loanwords in Hungarian cannot serve as a testimony of Old Slavic accent shifts, but merely of the place of the original (pitch) accent; and that the Slavic language that provided loanwords in early Hungarian must have been fairly uniform. As for Latin, it likewise renders an appeal to late accent shifts unnecessary.

Helimski also showed that a subset of Latin words containing a medial cluster *\*-CiV-* could trigger front-vowel harmony even if the original accent fell on a back vowel. Here it is shown that the distribution of front and back vocalism in this type is further governed by the vowel of the initial syllable. This minor rule possibly applies to Slavic as well.

## **1 Hungarian vowel harmony and Helimski's rule**

Hungarian, like many other Uralic languages, possesses vowel harmony, which is usually defined as a system of progressive assimilation by which vowels in the first syllable of a word dictates the vowel quality in suffixal syllables (inflectional endings, derivational endings etc.). From a diachronic viewpoint, since the harmony does not only affect suffixes, it

is more correct to say that any given word, derived or underived, must contain vowels from a certain restricted group, and that it is not necessarily a progressive assimilation – it may have started out as regressive, or even both at the same time, if the trigger occurred in the middle of the word. Vowel harmonic rules vary from language to language. In Hungarian, all native words must contain either back vowels (*a, o, u, á, ó, ú*) or “intolerant” front vowels (*e, ü, ö, ő, ő*) while they can always contain “tolerant” front vowels (*i, í, ë and é*).

In a brilliant and important article, Helimski (1992) showed that original stress in the lending language determines whether loanwords in Hungarian acquire front (“palatal”) vocalism or back (“velar”) vocalism. As he mentions, there has been no issue regarding the many Turkic loanwords because they were already subject to a similar vowel harmony from the outset (and in general consist of vowels that form part of the Hungarian phonemic system). However, Hungarian also absorbed many loanwords from languages without vowel harmony since the Magyar migrations into Central Europe in the 9<sup>th</sup> century, especially (unspecified) Slavic, Latin and different stages of High German, all of which have exported numerous polysyllabic words with back-vowel syllables interchanging with front-vowel syllables into Hungarian. Although in a few cases Hungarian ended up with both a back-vowel and a front-vowel variant, most often there is only one form.

Helimski showed that the outcome was determined by the original vocalism of syllable carrying the main stress in any of the three source languages. The rule is completely consistent for High German loanwords. In this article, I will try to explain a number of loanwords from Latin and Slavic that seem to deviate from the rule.

## 2 Loanwords from German

In loanwords from Old High German, Middle High German and Early Modern High German, the stressed syllable was normally also the first syllable; however, as Helimski pointed out (1998: 46), this was not *always* the case. The dialectal Hungarian word *isptály* ‘hospital’, borrowed from Middle High German, is important because it was itself a loanword in Middle High German where it had retained the original stress on the last syllable (as still in Modern High German *Spi<sup>l</sup>tal*):

- Hu. *bajor* 'Bavarian' (11th c. *Paiur*) < MHG *bayer* (Mollay 1982 : 150-151)  
 Hu. *cukor* 'sugar' (1587 : *cukor*) < Early NHG *czukcher*, *zugher* (Mollay 1982 : 150-151)  
 Hu. (obs.) *föröstök* (about 1395 : *felefstekum*) 'breakfast' < Early NHG *Früstukch*, *fruestukh* (Mollay 1982 : 278-289)  
 Hu. *herceg* 'prince' (1201: *Herceg*) < OHG *herzog* (Mollay 1982: 308)  
 Hu. dial. *isptály* (Jókai Kódex 1372/1448 *hyfpital*-, *Spital*-, 1527 *hyftaly*) < MHG *spital* (cf. NHG *Spi'tal*)  
 Hu. *kalmár* (1301: *Kalamar*-) < MHG *krāmære* 'tradesman, merchant' (Mollay 1982 : 336)  
 Hu. (obs.) *lőböstök* (1604 *Lo<sup>e</sup>bo<sup>e</sup>sto<sup>e</sup>c*) 'the plant Levisticum' < Early NHG *lūbestock*, *lebestock* (< Lat. *levisticum*) (Mollay 1982 : 389)  
 Hu. *polgár* 'citizen' (1229/1550 : *Pulgar*) < MHG *burgære* 'city-dweller' (Mollay 1982 : 444-446)  
 Hu. *püspök* 'bishop' (1177/about 1500 *Pyspek*- ; dial. *pöspök*, *pözsök*, *püspök*, *pispök*) < Old Bavarian *piscof* (Mollay 1982 : 459-465)  
 Hu. *tenyér* 'palm' < OHG *tenar*, MHG *tener* (Mollay 1982 : 528-530)  
 Hu. *vánkos* 'pillow' < MHG *wangechusse*, *wangküss*

### 3 Loanwords from Latin

Helinski further showed that in Latin loanwords the same rule applied, however only in such loanwords where vocalic harmonization was not blocked because of their bookish character and influence from the liturgical tradition. The affected words comprise many proper names. The following heterovocalic Latin words acquire back-vowel harmony:

- Hu. PN *Adorján* < Lat. *Hadriánus* (Fludorovits 1930: 17, 44)  
 Hu. *angyal* 'angel' < Lat. *ángelus* (TESz)  
 Hu. PN *Damján* < Lat. *Damiánus* (Fludorovits 1930: 17, 44 ; see below on the variant *Dëmjén*)  
 Hu. *kamara* 'chamber', *kamra* 'storeroom' < Lat. *cámara* (Fludorovits 1937: 41, TESz)<sup>1</sup>  
 Hu. *kanonok* (about 1405: *kananok*) 'clergyman' < Lat. *canónicus* (TESz)

<sup>1</sup> This word was probably borrowed via OHG *kamara*, *kamera* (Mollay 1982: 337-339)

Hu. PN *ibolya* (1340 *iwola*) 'violet' < Vulg.Lat., LLat. *vióla* < Class.Lat. *viola* (TESz)

Heterovocalic Latin words that acquire front-vowel harmony are more numerous. They comprise words with original stress on a front-vowel:

- Hu. PN *Elek* (1219/1550 *Elexius*) ← Lat. *Aléxius* (Fludorovits 1930: 9)  
 Hu. PN *Erzsébet* ← Lat. (?) *Elísabeth* or *Elisabéth* (Fludorovits 1930: 9)  
 Hu. PN *Ferenc* ← shortened form of Lat. *Franciscus* (Fludorovits 1930: 9)  
 Hu. *fülemüle*, *fülemile* (1395 *filemýle*) 'nightingale' ← Lat. *philoméla* (TESz)  
 Hu. PN *Lőrinc* (1138/1329 *Leurenci*) ← Lat. *Lauréntius* (Fludorovits 1930: 12)  
 Hu. *mise* (1372/1448 *mifet* acc.sg.) 'mass' ← Lat. *míssa* (TESz)  
 Hu. *petrezselyem* 'parsley' ← Lat. *petrosílium* (added by me; cf. Benkő & Imre 1972: 187)  
 Hu. *sekréstye* (1510 *sekreftye*-) 'sacristy' ← Lat. *sacrístia* (TESz)  
 Hu. *szerecsén* (1138/1329 *Scerecín*) 'Saracen' ← Lat. *Saracénus* (TESz)  
 Hu. *zsöllye* (about 1510: *sellye*) 'armchair (obs.); 'stall (in a theatre), orchestra seat, parterre seat' ← Lat. *sélla* 'seat, chair' (TESz)

Latin words with a medial cluster containing a palatal element in it, i.e. words containing a medial structure -CiV-, occur mostly with front vowels even if the originally stressed vowel (and many other vowels of the word) is a back vowel. In such cases variation sometimes occurs, and in the case of *Adorján* mentioned above no front-vowel variant is known:

- Hu. PN *Adorján* ← Lat. *Hadriánus* (Fludorovits 1930: 17, 44)  
 Hu. PN *Cerjék*, *Cirjék* ← Lat. *Cyriácus* (Fludorovits 1930: 7)  
 Hu. PN *Dëmjén*, variant of *Damján* ← Lat. *Damiánus* (Fludorovits 1930: 17, 44)  
 Hu. PN *Sebestyén* ← Lat. *Sebastiánus* (Fludorovits 1930: 7)

Helimski does not ascribe the variation in these four individual cases to any kind of regularity. It is noteworthy, however, that the first vowel in both *Hadriánus* and *Damiánus* are back-vowels, leading to back vocalism (*Adorján*, *Damján*) with a front-vowel variant in one of the cases (*Dëmjén*), while the first vowel of *Cyriácus* and *Sebastiánus* are front

vowels, leading to front-vowel forms only (*Cerjék/Cirjék*, *Sebestyén*). Although the material is limited we may tentatively assume that if a Latin word contained medial consonantal *-i-*, it will always end up having front vocalism in Hungarian (regardless of the quality of the stressed vowel) *unless both* the original first vowel and the original stressed vowel are back vowels.

Helimski ran into two real exceptions which exhibit back vocalism although the stressed vowel in Latin was a back vowel<sup>2</sup>:

Hu. *dézsma* ‘tithe’ ← Lat. *décima*

Hu. *tégla* ‘brick’ ← Lat. *tégula*

He suggests (p. 50) that “the high degree of compatibility of the harmonically tolerant vowel *é* with the back vowels overweighed its palatal character”. While these two words are indeed the only examples containing an open-syllable *-e-* in Latin, Helimski does not formulate his suggestion as a rule, in fact he exactly counts these examples as exceptions; and it is still a bit difficult to account for the fact that they did not end up as †*dézsme* or †*tégle* respectively. Below we will see how another rule must have applied which can also explain similar exceptions within the bulk of Old Slavic loanwords.

#### 4 Loanwords from Slavic

Slavic loanwords into Old Hungarian poses difficulties to a higher extent. First of all, we do not know all details of the exact language which was the main or perhaps even the only provider of Slavic loanwords, in all probability the Pannonian dialect of Late Common Slavic (Helimski 50). It is therefore crucial that we can find clear patterns in the material. Helimski shows that by using the same indicator, the accented syllable of the Slavic word, which is generally preserved in Russian (though as stress rather than pitch accent), we can see that, again, the place of ar-

<sup>2</sup> As he notes, other exceptions are actually not direct borrowings from Latin, but have been taken over from Old High German, as in the case of *monostor* ‘monastery’ where the immediate source was OHG *munusturi*, *munsturi* and Latin *monastĕrium* only the ultimate source. This is most likely also the case of several proper names like *Márton* (< *Mártin* rather than Lat. *Martínus*) and *Ágoston* (< *Águstín*, not Lat. *Augustínus*).

ticulation of the stressed vowel in the source language governs the vowel harmony of the Hungarian word:

- Hu. *acél* 'steel' (not †*ecél*) ← LCS \**ǫcēlb*; Kniesza 1974: 59-60)  
 Hu. *csahol* 'long shirt' (not †*csehöl*) ← LCS \**čexǫlb*, cf. Ru. *čexól* 'cover')  
 Hu. *csata* 'fight; (obs.) military team' (not †*csete*) ← LCS \**četa*, cf. Ru. *četá* 'match, pair'  
 Hu. *csónak* 'boat' (not †*csőnek*) ← LCS \**čǫlnǫkb*, cf. Ru. *čelnók* 'small boat'  
 Hu. *csorda* 'herd' (not †*csérde*) ← LCS \**čerda* > *čřěda*, cf. Ru. *čeredá* 'turn, order'  
 Hu. *csütörtök* 'Thursday' (not †*csutortok*) ← LCS *četvǫrtǫkb*, cf. ORu. *četvǫrtǫkb* (changed in Mod.Ru. *četvertók* 'four pieces')  
 Hu.dial. *deget* 'wheel-grease' (not e.g. †*dagat*) ← LCS \**degǫtb*, cf. Ru. *děgot* 'tar, pitch'  
 Hu. *ebéd* 'dinner' (not †*abéd*) ← LCS \**oběd* (Kniesza 1974: 166-167)  
 Hu. *kalapál* 'to hammer' (not †*kelepél*) ← LCS \**klepa-*, cf. Ru. *klepát* 'to rivet'  
 Hu. *morotva* 'old riverbed' (not †*mörötve*) ← LCS \**mǫrtva*, cf. Ru. *mertvá* 'dead (fem.)'  
 Hu. *öszton* 'stimulus, drive' (not †*osztón*) ← LCS \**ostǫnb*, cf. Ru. *ostén id.*  
 Hu. *rosta* 'sifter, sieve' (not e.g. †*restő*) ← LCS \**rešeto*, cf. Ru. *rešetó id.*  
 Hu. *szalonna* 'lard, bacon' (not e.g. †*szelönne*) ← LCS \**solnīná* > \**slanīná* '(salted) lard' (Kniesza 1974: 487)<sup>3</sup>  
 Hu. *szelénce* 'casket, small box' (not e.g. †*szalonca*) ← LCS \**solbnīca* 'box for salt, salt-cellar' (Kniesza 1974: 495-497)  
 Hu. *szërda* 'Wednesday' (not e.g. †*szarda*) ← LCS \**serda* > \**srěda*, cf. Ru. *sredá id.*  
 Hu. dial. (and obs.) *szosztra* ~ *szesztra* 'junior nun' (not e.g. †*szösztre* or †*szesztre*) < LCS \**sestra*, cf. Ru. *sestrá* 'sister'

Helimski finds the following exceptions which he notes might be explained by Pannonian accent shifts shared by Kaikavian Serbo-Croatian dialects of Slavonia (a suggestion by Sergej Nikolaev and G. Zamjatina):

<sup>3</sup> The same Slavic word entered Hungarian later via Romanian *slanină*, when the adaptation rules had stopped working, and became *szlanina* (Bakos 1977: 132).

Hu. *járom* ‘yoke’ ← LCS \**jaróm̃*, cf. Ru. *jarëm* id.  
 Hu. *szilva* ‘plum’ ← LCS \**slīva*, cf. Ru. *slīva* id.

There are also harmonic doublets of Slavic origin which are tentatively ascribed to accentual doublets in Slavic:

Hu. *csalad* ‘family’ ~ *cseled* ‘domestic, servant’ ← LCS \**čeljadb*  
 Hu. *ke|ř
|  |*c ~ dial. *katroc* ‘cage’ ← LCS \**kotar̃cb*  
 Hu. *varsa* ~ dial. *verse*, *vörse* ‘fishweir’ ← LCS \**vbrša*

respectively. Below we will see how another rule must have applied which can also explain similar exceptions within the bulk of Old Slavic loanwords.

## 5 Explaining the exceptions

The exceptions all contain a labial sound following the original stress which could have triggered a strategy for nativization dominated by back-vocalism. Perhaps Latin *-l-* also had this effect since Lat. *-u-* is likely to represent a centralized epenthetic vowel:

Hu. *dézsma* ‘tithe’ ← Lat. *décima*  
 Hu. *tégla* ‘brick’ ← Lat. *tégula*  
 Hu. *járom* ‘yoke’ ← LCS \**jaróm̃*, cf. Ru. *jarëm* id.  
 Hu. *szilva* ‘plum’ ← LCS \**slīva*, cf. Ru. *slīva* id.

At the same time, none of the regular examples contained an original labial. This may then also provide an alternative explanation for

Hu. PN *ibolya* (1340 *iwola*) ‘violet’ < Vulg.Lat., LLat. *vióla* < Class.Lat. *Víola*

where we do not have to appeal to a Late Latin or Vulgar Latin pronunciation, even if it did apply, because our rule works perfectly even with the Classical Latin pronunciation.

The rule seems to be exceptionless. In Helimski’s material I can find the following other words with a (non-initial) labial, one from Slavic and one from Latin ; in both cases you have front-vocalism because the



source word had an accented front vowel, and this is allowed because the labial sequence precedes, not follows, the accented vowel.

Hu. *ebéd* 'dinner' (not †*abéd*) ← LCS \**oběd*

Hu. *fülemüle*, *fülemile* (1395 *filemýle*) 'nightingale' ← Lat. *philomēla*

Another example might be added, namely

Hu. *bazsalikom* (not e.g. †*bezselikem*) ← Lat. *basílicum*

where front-vowel harmony, which should otherwise occur because of the Latin stressed front-vowel *-i-*, is blocked by virtue of the labial sequence *-um*<sup>4</sup>. Another herb name, not mentioned in Helimski's article either, is

Hu. *petrezselyem* 'parsley' (not †*patrazsalyam*) ← Lat. *petrosílium*

The latter does have front-vowel harmony throughout the word despite the labial element—because here our other rule of medial palatal sequences *-CiV-* triggers it. The two words are further distinct by their place of vowel articulation in the first syllable which may have played a role at least for *bazsalikom* in the sense that it could have contributed to add up the factors in favor of a back vowel.

With this rule we can even get solved where the stress was in the Latin form of *Elisabeth*; it is now clear that it must have been on the last syllable because it could not have preceded the labial *-b-* which would have resulted in a different vowel harmony with at least one back vowel substituting the *-a-*. This may even be applicable if the name was borrowed via another language than Latin. Some Slavic languages have stress on *-bét-* (e.g. Serbian *jelisavéta*) but the Hungarian form shows no trace of the feminine ending *-a* that has been added to the name. Thus, in a revised version:

Hu. PN *Erzsébet* ← Lat. (?) *Elisabéth* (not † *Elisabeth*)

<sup>4</sup> This example was suggested to me by Seán Vrieland.

## 6 Conclusion

I have tried to demonstrate that the accented vowel of a word in the source language governs whether it ends up in Hungarian with front-vowel or (predominant) back-vowel harmony to an even more regular extent than discovered by Helimski. He inferred that exceptions and variation within the Slavic material may be due to accent shifts reminiscent of those known from neighbouring Kajkavian dialects and is a potentially useful source for research in Balto-Slavonic accentology. As for the Latin material, he ascribed exceptions to the back-vowel tolerance of the neutral front-vowel *é* in Hungarian itself, and to accent shifts from Classical to Late and Vulgar Latin.

I hope to have shown that all these exceptions can be explained by a rather simple conditional rule of adaptation: If the accented vowel of the source language, whether (Pannonian) Slavic or Latin, was followed by a labial sound and/or *-l-*, front-vowel harmonization was blocked, and original vowels were substituted with back vowels and neutral front-vowels only. Although it may disappoint Slavicists who could wish for Hungarian as yet another testimony of old Slavic accentuation developments, I find it much more economical to embrace a solution that explains the Slavic and Latin irregularities by the same exceptionless rule.

With less certainty, we may also set up adaptation rules for a subset of Latin words (incidentally, all proper names) with a medial palatal cluster *-CiV-*. Helimski showed that this type mostly turns out with Hungarian front-vowel harmony even if the originally stressed vowel (and many other vowels in the word) was a back vowel. Again, he had to deal with variation and exceptions, but he simply deduced that vacillation between front-vowel harmony and back-vowel harmony was the result of two conflicting tendencies pulling in different directions. I pointed out that, although based on scarce material, there are at least no exceptions to a rule by which Latin words with a palatal medial cluster *-CiV-* acquires back-vocalism only if both the stressed vowel and the first vowel of the source word has back vocalism. Without these two triggering factors working together, front-vowel harmony overrules the default law.

Finally, I have suggested that harmonic doublets in the Slavic material, although potentially reflecting Slavic accentual doublets, may partly reflect the same tendencies, cf. that e.g. *család* 'family' vs. *cseled* 'domes-

tic, servant' is the only example with a medial structure \*-CjV- that is similar to the Latin structure \*-CiV- causing variation above.

The rules can be formalized this way, where VH represents vowel harmony and SL stands for source language:

VH[+front] if SL has V[+stress][+front]

VH [+back] if SL has V[+stress][+back]

## References

- Andersen, Lloyd B., 1980: "Using asymmetrical and gradient data in the study of vowel harmony". – Robert M. Vago: *Issues in Vowel Harmony. Proceedings of the CUNY Linguistics Conference on Vowel Harmony, 14th May 1977* [= *Studies in Language Companion Series (SLCS)* 6]. Amsterdam: John Benjamins.
- Bakos, Ferenc, 1977: "Les éléments roumains du lexique hongrois et quelques problèmes de l'emprunt linguistique". – *Acta Linguistica Academiae Scientiarum Hungaricae* 27 (1-2): 111-160.
- Benkő, Loránd & Samu Imre, 1972: *The Hungarian Language* [= *Janua Linguarum Series Practica* 134]. The Hague / Paris: Mouton.
- Fludorovits, Jolán, 1930: *Latin jövevényszavaink hangtana* [= *A Magyar Nyelvtudományi Társaság kiadványai* 26]. Budapest.
- Greenberg, Marc L., 2004: "Sifting the evidence for the reconstruction of Pannonian Slavic". Review of Richards 2003. – *Canadian Slavonic Papers / Revue Canadienne des Slavistes* 43,1/2: 213-220.
- Helinski, Eugen, 2003: "Slavic/Latin/German Stress and Hungarian vowel Harmony". – László Honti, Sirkka-Liisa Hahmo, Tette Hofstra, Jolanta Jastrzębska & Osmo Nikkilä (ed.): *Finnisch-ugrische Sprachen zwischen dem germanischen und dem slavischen Sprachraum*. Amsterdam / Atlanta: Rodopi. Pp. 45-54.
- Kniezsa, István, 1974: *A magyar nyelv szláv jövevényszavai I*. Budapest: Akadémiai Kiadó.
- Laakso, Johanna. 2006. Review of Richards 2003. *Finnisch-Ugrische Forschungen* 56: 254.
- Mollay, Károly, 1982: *Német-Magyar Nyelvi Érintkezések a XVI. Század Végeig*. Budapest: Akadémiai Kiadó.
- Nuorloto, Juhani, 2010: "Central Slovak and Kajkavian Structural Convergences: A Tentative Survey". *Slovo* 50: 37-45.
- Richards, Ronald O., 2003: *The Pannonian Slavic Dialect of the Common Slavic Protolanguage: The View from Old Hungarian* [= *UCLA Indo-European Studies*, vol. 2]. Los Angeles.
- TESz = *A Magyar nyelv történeti-etimológiai szótára I-III*. Budapest 1967-1976.
- Vishogradskaja, Ina: "Vowel harmony in the early borrowings from Slavic to Hungarian: Proof of root vowel harmony?" Ms.

## English Summary

### ***Word Exchange at the Gates of Europe: Five Millennia of Language Contact***

Indo-European and Uralic languages dominate present-day Europe, but both families are newcomers which replaced most of the indigenous languages step by step from the Bronze Age onwards. The encounter between indigenous and intrusive cultures, however, was most certainly not the only interaction that took place. By the time of arrival in Europe, the Indo-European and Uralic populations had already broken up and constituted a patchwork of languages and cultures that continued the process of convergence and interchange. Whether contacts were connected to trade, war, social interaction, or exchange of inventions is revealed by the character of the loanwords in each individual case – while the shape of the loanwords expose the time depth and the direction of borrowing.

Traditionally, scholars have thought that basically all loanwords between Indo-European and Uralic languages went in one direction – from the former to the latter. Such an asymmetry is supposed to reflect a past relationship between two peoples where one had the upper hand, technically and politically, at the time of borrowing.

In this dissertation it is shown that cultures of the Northeast played a surprisingly important role in the shaping of our continent from prehistoric to Medieval times; and it is shown how these circumstances are reflected even in the vocabularies of modern European languages.

The Indo-European tribes, shortly after their migrations into Europe, came to form part of new cultural communities, influenced by Uralic populations from the North. This had a significant impact on specific parts of the vocabulary, notably terms for religion and warfare. Many trade terms (such as Danish *pung* ‘purse’), and words for tools (e.g. *hammer*) and religious concepts (e.g. *hell*) originate from Fenno-Ugric

and other languages spoken in Northeastern Europe at the time. Even our word *half* can be shown to derive from an old Fennic trading term meaning ‘reduced, cheap (of prices)’.

Some terms denote animals hunted for their pelt (e.g. *mink*) and were exchanged in connection to centuries of fur trade along the Baltic coasts from the Roman Ages to the Hanseatic period. Other words for animals, among them quite a few used for pigs and boars, are, quite astonishingly, much older loans going back the pitted-ware culture around 3000 BC. Some loanwords show, for the first time, that (Proto-)Celts and Fennic peoples must have been in direct contact with each other.

## Dansk Resumé

### ***Ordudveksling ved Europas porte: Fem årtusinders sprogkontakt***

Indoeuropæiske og uralske sprog dominerer i dag Europa, men begge sprogfamilier er historisk set nytilkomne, som skridt for skridt erstattede Europas oprindelige sprog fra bronzealderen og frem. Mødet mellem de indfødte og indtrængende kulturer var dog helt sikkert ikke den eneste interaktion, der fandt sted. Da indoeuropæisk- og uralsktalende stammer nåede til Østeuropa, var de allerede splittet op og udgjorde et kludetæppe af sprog og kulturer, som vedblev at være i kontakt og påvirke hinanden på kryds og tværs. Hvilke historiske situationer, de enkelte sprogkontakter reflekterer – om der var tale om handel, krig, social interaktion eller udveksling af opfindelser – afsløres af låneordenes karakter. Samtidig kan deres form afsløre, hvornår og i hvilken retning lånet fandt sted.

Traditionelt har sprogforskerne ment, at nærmest alle låneord mellem indoeuropæiske og uralske sprog var lånt netop *fra* indoeuropæisk til uralsk – stort set aldrig den anden vej. Man regner med, at sådanne asymmetrier afspejler fortidige relationer, hvor folkeslaget med det långivende sprog var teknologisk og politisk overlegent på lånetidspunktet.

I denne afhandling påvises det, at nordøstlige kulturer har spillet en overraskende vigtig rolle for udformningen af vores kontinent fra forhistorisk tid til langt ind i middelalderen, og hvordan disse forhold stadig afspejles i moderne europæiske sprogs ordforråd.

Kort tid efter at de indoeuropæiske stammer ankom til Europa blev de hver især en del af nye kulturelle fællesskaber, påvirket af uralske folkeslag nordpå. Dette fik stor indflydelse på bestemte dele af ordforrådet, især termer inden for religion og krigsførelse. Mange handelstermer (fx *pung*), ord for redskaber (fx *hammer*) og religiøse termer (fx *helvede*) stammer fra finsk-ugrisk og andre sproggrupper fra datidens nordøst.

Selv vores ord *halv* kan påvises at gå tilbage til en gammel finsk handels-term, der betyder ‘reduceret, billig’.

Nogle af termene har med pelsdyr at gøre (fx *mink*) og er udvekslet i forbindelse med den pelshandel, der foregik langs Østersøens kyster i århundreder fra romertid til hansetid. Andre dyretermer, fx en lang række ord for grise og svin, er overraskende meget ældre lån, der har forbindelse måske helt tilbage til den grubekeramiske kultur omkring 3000 f.Kr. Nogle låneord viser for første gang, at (ur)keltere og finske folk må have været i direkte kontakt med hinanden.

